PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 7 :
C12N 15/12, 5/10, 1/21, C07K 14/52, 16/22, G01N 33/50, 33/566, A01K 67/027, A61K 31/70, 38/18, 39/395, A61P 9/00

A2

(11) International Publication Number: WO 00/37641

(43) International Publication Date: 29 June 2000 (29.06.00)

(21) International Application Number: PCT/US99/30503

(22) International Filing Date: 21 December 1999 (21.12.99)

(30) Priority Data:

9828377.3	22 December 1998 (22.12.98)	GB
60/124,967	18 March 1999 (18.03.99)	US
60/164,131	8 November 1999 (08.11.99)	US

(71) Applicant *(for all designated States except US)*: JANSSEN PHARMACEUTICA N.V. [BE/BE]; Turnhoutseweg 30, B–2340 Beerse (BE).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: GORDON, Robert, Douglas [GB/BE]; Janssen Pharmaceutica N.V., Turnhoutseweg 30, B–2340 Beerse (BE). SPRENGEL, Jorg, Jurgen [DE/BE]; Janssen Pharmaceutica N.V., Turnhoutseweg 30, B–2340 Beerse (BE). YON, Jeffrey, Roland [GB/BE]; Janssen Pharmaceutica N.V., Turnhoutseweg 30, B–2340 Beerse (BE). DIJKMANS, Josiena, Johanna, Huberdina [NL/BE]; Janssen Pharmaceutica N.V., Turnhoutseweg 30, B–2340 Beerse (BE). GOSIEWSKA, Anna [PL/US]; Johnson & Johnson Consumer Products, Wound Healing Technology Resource Center, RG24, North Building, 199 Grandview Road, Skillman, NJ 08558 (US). DHANARAJ, Sridevi, Naidu [IN/US]; Johnson & Johnson Consumer Products, Wound Healing Technology Resource Center, RG24, North Building, 199 Grandview Road, Skillman, NJ 08558 (US). XU, Jean [CN/US]; Johnson & Johnson Consumer Products, Wound Healing Technology Resource Center, RG24, North Building, 199 Grandview Road, Skillman, NJ 08558 (US).

(74) Agent: VAN AMSTERDAM, John, R.; Wolf, Greenfield & Sacks, P.C., 600 Atlantic Avenue, Boston, MA 02210 (US).

(81) Designated States: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

Without international search report and to be republished upon receipt of that report.

With an indication in relation to deposited biological material furnished under Rule 13bis separately from the description.

(54) Title: VASCULAR ENDOTHELIAL GROWTH FACTOR–X

(57) Abstract

There is provided a novel vascular endothelial growth factor, herein designated VEGF–X, in addition to the nucleic acid molecule encoding it, a host cell transformed with said vector and compounds which inhibit or enhance angiogenesis. Also provided is the sequence of a CUB domain present in the sequence of VEGF–X which domain itself prevents angiogenesis and which is used to treat diseases associated with inappropriate vascularisation or angiogenesis.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

VASCULAR ENDOTHELIAL GROWTH FACTOR-X

The present invention is concerned with a novel vascular endothelial growth factor (VEGF) herein designated "VEGF-X", and characterisation of the nucleic acid and amino acid sequences of VEGF-X.

Introduction

Angiogenesis involves formation and proliferation of new blood vessels, and is an essential physiological process for normal growth and development of tissues in, for example, embryonic development, tissue regeneration and organ and tissue repair. Angiogenesis also features in the growth of human cancers which require continuous stimulation of blood vessel growth. Abnormal angiogenesis is associated with other diseases such as rheumatoid arthritis psoriasis and diabetic retinopathy.

Capillary vessels consist of endothelial cells which carry the genetic information necessary to proliferate to form capillary networks. Angiogenic molecules which can initiate this process have previously been characterised. A highly selective mitogen for vascular enothelial cells is vascular endothelial growth factor (VEGF) (Ferrara *et al.*, "Vascular Endothelial Growth Factor: Basic Biology and Clinical Implications". Regulation of angiogenesis, by I.D. Goldberg and E.M. Rosen 1997 Birkhauser Verlag Basle/Switzerland). VEGF is a potent vasoactive protein which is comprised of a glycosylated cationic 46-49 kd dimer having two 24 kd subunits. It is inactivated by sulfhydryl reducing agents and is resistant to acidic pH and to heating and binds to immobilised heparin.

VEGF-A has four different forms of 121, 165, 189 and 206 amino acids respectively due to alternative splicing. VEGF121 and VEGF165 are soluble and are capable of promoting angiogenesis, whereas VEGF189 and VEGF206 are bound to heparin containing proteoglycans in the cell surface. The temporal and spatial expression of VEGF has been correlated with physiological proliferation of the blood vessels (Gajdusek, C.M., and Carbon, S.J., Cell Physiol., 139:570-579, (1989)); McNeil, P.L., Muthukrishnan, L., Warder, E., D'Amore, P.A., J. Cell. Biol., 109:811-822, (1989)). Its high affinity binding sites are localized only on endothelial cells in tissue sections (Jakeman, L.B., *et al.*, Clin. Invest. 89:244-253 (1989)). The growth factor can be isolated from pituitary cells and several tumor cell lines, and has been implicated in some human gliomas (Plate, K.H. Nature 359:845-848, (1992)). The inhibition of VEGF function by anti-VEGF monoclonal antibodies was shown to inhibit tumor growth in immune-deficient mice (Kim, K.J., Nature 362:841-844, (1993)).

VEGF proteins have been described in the following patents and applications all of which are hereby incorporated by reference EP-0,506,477, WO-95/24473, WO-98/28621, WO-90/13649, EP-0,476,983, EP-0,550,296, WO-90/13649, WO-96/26736, WO-96/27007, WO-98/49300, WO-98/36075, WO-98/840124, WO-90/11084, WO-98/24811, WO-98/10071, WO-98/07832, WO-98/02543, WO-97/05250, WO-91/02058, WO-96/39421, WO-96/39515, WO-98/16551.

The present inventors have now identified a further vascular endothelial growth factor, designated herein as "VEGF-X", and the nucleic acid sequence encoding it, which has potentially significant benefits for the treatment of tumours and other conditions mediated by inappropriate angiogenic activity.

Summary of the Invention

In the present application, there is provided a novel vascular endothelial growth factor, herein designated "VEGF-X", nucleic acid molecules encoding said growth factor, an expression vector comprising said nucleic acid molecule, a host cell transformed with said vector and compounds which inhibit or enhance angiogenesis. Also provided is the sequence of a CUB domain present in the sequence of VEGF-X which domain itself prevents angiogenesis and which is used to treat diseases associated with inappropriate vascularisation or angiogenesis.

Detailed Description of the Invention

Therefore, according to a first aspect of the present invention there is provided a nucleic acid molecule encoding a VEGF-X protein or a functional equivalent, fragment, derivative or bioprecursor thereof, said protein comprising the amino acid sequence from position 23 to 345 of the amino acid sequence illustrated in Figure 10. Alternatively, the nucleic acid molecule of the invention encodes the complete sequence identified in Figure 10 and which advantageously includes a signal peptide to express said protein extracellularly. Preferably, the nucleic acid molecule is a DNA and even more preferably a cDNA molecule. Preferably, the nucleic acid molecule comprises the nucleotide sequence from position 257 to 1291 of the nucleotide sequence illustrated in Figure 9. In a preferred embodiment the nucleic acid is of mammalian origin and even more preferably of human origin.

In accordance with the present invention a functional

equivalent should be taken to mean a protein, or a sequence of amino acids that have similar function to the VEGF-X protein of the invention.

Also provided by this aspect of the present invention is a nucleic acid molecule such as an antisense molecule capable of hybridising to the nucleic acid molecules according to the invention under high stringency conditions, which conditions would be well known to those skilled in the art.

Stringency of hybridisation as used herein refers to conditions under which polynucleic acids are stable. The stability of hybrids is reflected in the melting temperature (Tm) of the hybrids. Tm can be approximated by the formula:

$$81.5°C+16.6(\log_{10}[Na^+]+0.41\ (\%G\&C)-600/l$$

wherein l is the length of the hybrids in nucleotides. Tm decreases approximately by 1-1.5°C with every 1% decrease in sequence homology.

The term "stringency" refers to the hybridisation conditions wherein a single-stranded nucleic acid joins with a complementary strand when the purine or pyrimidine bases therein pair with their corresponding base by hydrogen bonding. High stringency conditions favour homologous base pairing whereas low stringency conditions favour non-homologous base pairing.

"Low stringency" conditions comprise, for example, a temperature of about 37°C or less, a formamide concentration of less than about 50%, and a moderate to low salt (SSC) concentration; or, alternatively, a temperature of about 50°C or less, and a moderate to high salt (SSPE) concentration, for example 1M NaCl.

"High stringency" conditions comprise, for example, a temperature of about 42°C or less, a formamide concentration of less than about 20%, and a low salt (SSC) concentration; or, alternatively, a temperature of about 65°C, or less, and a low salt (SSPE) concentration. For example, high stringency conditions comprise hybridization in 0.5 M $NaHPO_4$, 7% sodium dodecyl sulfate (SDS), 1 mM EDTA at 65°C (Ausubel, F.M. *et al.* Current Protocols in Molecular Biology, Vol. I, 1989; Green Inc. New York, at 2.10.3).

"SSC" comprises a hybridization and wash solution. A stock 20X SSC solution contains 3M sodium chloride, 0.3M sodium citrate, pH 7.0.

"SSPE" comprises a hybridization and wash solution. A 1X SSPE solution contains 180 mM NaCl, 9mM Na_2HPO_4 and 1 mM EDTA, pH 7.4.

The nucleic acid capable of hybridising to nucleic acid molecules according to the invention will generally be at least 70%, preferably at least 80 or 90% and more preferably at least 95% homologous to the nucleotide sequences according to the invention.

The antisense molecule capable of hybridising to the nucleic acid according to the invention may be used as a probe or as a medicament or may be included in a pharmaceutical composition with a pharmaceutically acceptable carrier, diluent or excipient therefor.

The term "homologous" describes the relationship between different nucleic acid molecules or amino acid sequences wherein said sequences or molecules are related by partial identity or similarity at one or more blocks or regions within said molecules or

sequences.

The present invention also comprises within its scope proteins or polypeptides encoded by the nucleic acid molecules according to the invention or a functional equivalent, derivative or bioprecursor thereof.

Therefore, according to a further aspect of the present invention, there is provided a VEGF-X protein, or a functional equivalent, derivative or bioprecursor thereof, comprising an amino acid sequence from position 23 to 345 of the sequence as illustrated in Figure 10, or alternatively which amino acid sequence comprises the complete sequence of Figure 10. A further aspect of the invention comprises a VEGF-X protein, or a functional equivalent, derivative or bioprecusor thereof, encoded by a nucleic acid molecule according to the invention. Preferably, the VEGF-X protein encoded by said nucleic acid molecule comprises the sequence from position 23 to 345 of the amino acid sequence as illustrated in Figure 10, or which sequence alternatively comprises the sequence of amino acids of Figure 10.

The DNA molecules according to the invention may, advantageously, be included in a suitable expression vector to express VEGF-X encoded therefrom in a suitable host. Incorporation of cloned DNA into a suitable expression vector for subsequent transformation of said cell and subsequent selection of the transformed cells is well known to those skilled in the art as provided in Sambrook *et al.* (1989), molecular cloning, a laboratory manual, Cold Spring Harbour Laboratory Press.

An expression vector according to the invention includes a vector having a nucleic acid according to

the invention operably linked to regulatory sequences, such as promoter regions, that are capable of effecting expression of said DNA fragments. The term "operably linked" refers to a juxta position wherein the components described are in a relationship permitting them to function in their intended manner. Such vectors may be transformed into a suitable host cell to provide for expression of a polypeptide according to the invention. Thus, in a further aspect, the invention provides a process for preparing polypeptides according to the invention which comprises cultivating a host cell, transformed or transfected with an expression vector as described above under conditions to provide for expression by the vector of a coding sequence encoding the polypeptides, and recovering the expressed polypeptides.

The vectors may be, for example, plasmid, virus or phage vectors provided with an origin of replication, and optionally a promoter for the expression of said nucleotide and optionally a regulator of the promoter.

The vectors may contain one or more selectable markers, such as, for example, ampicillin resistance.

Regulatory elements required for expression include promoter sequences to bind RNA polymerase and transcription initiation sequences for ribosome binding. For example, a bacterial expression vector may include a promoter such as the lac promoter and for translation initiation the Shine-Dalgarno sequence and the start codon AUG. Similarly, a eukaryotic expression vector may include a heterologous or homologous promoter for RNA polymerase II, a downstream polyadenylation signal, the start codon AUG, and a termination codon for detachment of the

ribosome. Such vectors may be obtained commercially or assembled from the sequences described by methods well known in the art.

Nucleic acid molecules according to the invention may be inserted into the vectors described in an antisense orientation in order to provide for the production of antisense RNA. Antisense RNA or other antisense nucleic acids may be produced by synthetic means.

In accordance with the present invention, a defined nucleic acid includes not only the identical nucleic acid but also any minor base variations including in particular, substitutions in cases which result in a synonymous codon (a different codon specifying the same amino acid residue) due to the degenerate code in conservative amino acid substitutions. The term "nucleic acid sequence" also includes the complementary sequence to any single stranded sequence given regarding base variations.

The present invention also advantageously provides nucleic acid sequences of at least approximately 10 contiguous nucleotides of a nucleic acid according to the invention and preferably from 10 to 50 nucleotides even more preferably, the nucleic acid sequence comprise the sequences illustrated in Figure 3. These sequences may, advantageously be used as probes or primers to initiate replication, or the like. Such nucleic acid sequences may be produced according to techniques well known in the art, such as by recombinant or synthetic means. They may also be used in diagnostic kits or the like for detecting the presence of a nucleic acid according to the invention. These tests generally comprise contacting the probe with the sample under hybridising conditions and detecting for the presence of any duplex or triplex

formation between the probe and any nucleic acid in the sample.

The nucleic acid sequences according to this aspect of the present invention comprise the sequences of nucleotides illustrated in Figures 3 and 5.

According to the present invention these probes may be anchored to a solid support. Preferably, they are present on an array so that multiple probes can simultaneously hybridize to a single biological sample. The probes can be spotted onto the array or synthesised *in situ* on the array. (See Lockhart *et al.*, Nature Biotechnology, vol. 14, December 1996 "Expression monitoring by hybridisation to high density oligonucleotide arrays". A single array can contain more than 100, 500 or even 1,000 different probes in discrete locations.

The nucleic acid sequences, according to the invention may be produced using such recombinant or synthetic means, such as for example using PCR cloning mechanisms which generally involve making a pair of primers, which may be from approximately 10 to 50 nucleotides to a region of the gene which is desired to be cloned, bringing the primers into contact with mRNA, cDNA, or genomic DNA from a human cell, performing a polymerase chain reaction under conditions which brings about amplification of the desired region, isolating the amplified region or fragment and recovering the amplified DNA. Generally, such techniques are well known in the art, such as described in Sambrook *et al.* (Molecular Cloning: a Laboratory Manual, 1989).

The nucleic acids or oligonucleotides according to the invention may carry a revealing label. Suitable

labels include radioisotopes such as ^{32}P or ^{35}S, enzyme labels or other protein labels such as biotin or fluorescent markers. Such labels may be added to the nucleic acids or oligonucleotides of the invention and may be detected using known techniques *per se*.

Advantageously, human allelic variants or polymorphisms of the DNA molecule according to the invention may be identified by, for example, probing cDNA or genomic libraries from a range of individuals, for example, from different populations. Furthermore, nucleic acids and probes according to the invention may be used to sequence genomic DNA from patients using techniques well known in the art, such as the Sanger Dideoxy chain termination method, which may, advantageously, ascertain any predisposition of a patient to certain disorders associated with a growth factor according to the invention.

The protein according to the invention includes all possible amino acid variants encoded by the nucleic acid molecule according to the invention including a polypeptide encoded by said molecule and having conservative amino acid changes. Conservative amino acid substitution refers to a replacement of one or more amino acids in a protein as identified in Table 1. Proteins or polypeptides according to the invention further include variants of such sequences, including naturally occurring allelic variants which are substantially homologous to said proteins or polypeptides. In this context, substantial homology is regarded as a sequence which has at least 70%, preferably 80 or 90% and preferably 95% amino acid homology with the proteins or polypeptides encoded by the nucleic acid molecules according to the invention. The protein according to the invention may be recombinant, synthetic or naturally occurring, but is

preferably recombinant.

The nucleic acid or protein according to the invention may be used as a medicament or in the preparation of a medicament for treating cancer or other diseases or conditions associated with expression of VEGF-X protein.

Advantageously, the nucleic acid molecule or the protein according to the invention may be provided in a pharmaceutical composition together with a pharmacologically acceptable carrier, diluent or excipient therefor.

The present invention is further directed to inhibiting VEGF-X *in vivo* by the use of antisense technology. Antisense technology can be used to control gene expression through triple-helix formation of antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion or the mature DNA sequence, which encodes for the protein of the present invention, is used to design an antisense RNA oligonucleotide of from 10 to 50 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple-helix - see Lee *et al.* Nucl. Acids Res., 6:3073 (1979); Cooney *et al.*, Science, 241:456 (1988); and Dervan *et al.*, Science, 251: 1360 (1991), thereby preventing transcription and the production of VEGF-X. The antisense RNA oligonucleotide hybridises to the mRNA *in vivo* and blocks translation of an mRNA molecule into the VEGF-X protein (antisense - Okano, J. Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression, CRC Press, Boca Raton, FL (1988)).

Alternatively, the oligonucleotide described above can be delivered to cells by procedures in the art such that the anti-sense RNA and DNA may be expressed *in vivo* to inhibit production of VEGF-X in the manner described above.

Antisense constructs to VEGF-X, therefore, may inhibit the angiogenic activity of VEGF-X and prevent the further growth of or even regress solid tumours, since angiogenesis and neovascularization are essential steps in solid tumour growth. These antisense constructs may also be used to treat rheumatoid arthritis, psoriasis and diabetic retinopathy which are all characterized by abnormal angiogenesis.

A further aspect of the invention provides a host cell or organism, transformed or transfected with an expression vector according to the invention. The host cell or organism may advantageously be used in a method of producing VEGF-X, which comprises recovering any expressed VEGF-X from the host or organism transformed or transfected with the expression vector.

According to a further aspect of the invention there is also provided a transgenic cell, tissue or organism comprising a transgene capable of expressing VEGF-X protein according to the invention. The term "transgene capable of expression" as used herein means a suitable nucleic acid sequence which leads to expression of VEGF-X or proteins having the same function and/or activity. The transgene, may include, for example, genomic nucleic acid isolated from human cells or synthetic nucleic acid, including DNA integrated into the genome or in an extrachromosomal state. Preferably, the transgene comprises the nucleic acid sequence encoding the proteins according to the invention as described herein, or a functional

fragment of said nucleic acid. A functional fragment of said nucleic acid should be taken to mean a fragment of the gene comprising said nucleic acid coding for the proteins according to the invention or a functional equivalent, derivative or a non-functional derivative such as a dominant negative mutant, or bioprecursor of said proteins. For example, it would be readily apparent to persons skilled in the art that nucleotide substitutions or deletions may be used using routine techniques, which do not affect the protein sequence encoded by said nucleic acid, or which encode a functional protein according to the invention.

VEGF-X protein expressed by said transgenic cell, tissue or organism or a functional equivalent or bioprecursor of said protein also forms part of the present invention.

Antibodies to the protein or polypeptide of the present invention may, advantageously, be prepared by techniques which are known in the art. For example, polyclonal antibodies may be prepared by inoculating a host animal, such as a mouse or rabbit, with the polypeptide according to the invention or an epitope thereof and recovering immune serum. Monoclonal antibodies may be prepared according to known techniques such as described by Kohler R. and Milstein C., Nature (1975) 256, 495-497. Advantageously, such antibodies may be included in a kit for identifying VEGF-X in a sample, together with means for contacting the antibody with the sample.

Advantageously, the antibody according to the invention may also be used as a medicament or in the preparation of a medicament for treating tumours or other diseases associated with expression of VEGF-X.

The invention also further provides a pharmaceutical composition comprising said antibody together with a pharmaceutically acceptable carrier diluent or excipient therefor.

Proteins which interact with the polypeptide of the invention may be identified by investigating protein-interactions using the two-hybrid vector system first proposed by Chien *et al.*, (1991) Proc. Natl. Acad. Sci. USA 88 : 9578-9582.

This technique is based on functional reconstitution *in vivo* of a transcription factor which activates a reporter gene. More particularly the technique comprises providing an appropriate host cell with a DNA construct comprising a reporter gene under the control of a promoter regulated by a transcription factor having a DNA binding domain and an activating domain, expressing in the host cell a first hybrid DNA sequence encoding a first fusion of a fragment or all of a nucleic acid sequence according to the invention and either said DNA binding domain or said activating domain of the transcription factor, expressing in the host at least one second hybrid DNA sequence, such as a library or the like, encoding putative binding proteins to be investigated together with the DNA binding or activating domain of the transcription factor which is not incorporated in the first fusion; detecting any binding of the proteins to be investigated with a protein according to the invention by detecting for the presence of any reporter gene product in the host cell; optionally isolating second hybrid DNA sequences encoding the binding protein.

An example of such a technique utilises the GAL4 protein in yeast. GAL4 is a transcriptional activator of galactose metabolism in yeast and has a separate

domain for binding to activators upstream of the galactose metabolising genes as well as a protein binding domain. Nucleotide vectors may be constructed, one of which comprises the nucleotide residues encoding the DNA binding domain of GAL4. These binding domain residues may be fused to a known protein encoding sequence, such as for example, the nucleic acids according to the invention. The other vector comprises the residues encoding the protein binding domain of GAL4. These residues are fused to residues encoding a test protein. Any interaction between polypeptides encoded by the nucleic acid according to the invention and the protein to be tested leads to transcriptional activation of a reporter molecule in a GAL-4 transcription deficient yeast cell into which the vectors have been transformed. Preferably, a reporter molecule such as β-galactosidase is activated upon restoration of transcription of the yeast galactose metabolism genes.

A further aspect of the present invention also provides a method of identifying VEGF-X in a sample, which method comprises contacting said sample with an antibody according to the invention and monitoring for any binding of any proteins to said antibody. A kit for identifying the presence of VEGF-X in a sample is also provided comprising an antibody according to the invention and means for contacting said antibody with said sample.

VEGF-X may be recovered and purified from recombinant cell cultures by methods known in the art, including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxyapatite chromatography and lectin

chromatography.

The VEGF-X protein of the present invention may be a naturally purified product, or a product of chemical synthetic procedures, or produced by recombinant techniques from a prokaryotic or eukaryotic host (for example, by bacterial yeast, higher plant, insect and mammalian cells in culture). Depending upon the host employed in a recombinant production procedure, the polypeptides of the present invention may be glycosylated with mammalian or other eukaryotic carbohydrates or may be non-glycosylated.

VEGF-X is particularly advantageous as a wound healing agent, where, for example, it is necessary to re-vascularize damaged tissues, or where new capillary angiogenesis is important. Accordingly, VEGF-X may be used for treatment of various types of wounds such as for example, dermal ulcers, including pressure sores, venous ulcers, and diabetic ulcers. In addition, it can be used in the treatment of full-thickness burns and injuries where angiogenesis is desired to prepare the burn in injured sites for a skin graft and flap. In this case, VEGF-X or the nucleic acid encoding it may be applied directly to the wound. VEGF-X may be used in plastic surgery when reconstruction is required following a burn, other trauma, or even for cosmetic purposes.

An important application of VEGF-X is to induce the growth of damaged bone, periodontium or ligament tissue. For example, it may be used in periodontal disease where VEGF-X is applied to the roots of the diseased teeth, leading to the formation of new bone and cementum with collagen fibre ingrowths. It can be used for regenerating supporting tissues of teeth, including alveolar bone, cementum and periodontal

ligament, that have been damaged by disease and trauma.

Since angiogenesis is important in keeping wounds clean and non-infected, VEGF-X may be used in association with surgery and following the repair of cuts. It should be particularly useful in the treatment of abdominal wounds where there is a high risk of infection.

VEGF-X can also be used for the promotion of endothelialization in vascular graft surgery. In the case of vascular grafts using either transplanted or synthetic material, VEGF-X may be applied to the surface of the graft or at the junction to promote the growth of the vascular endothelial cells. One derivation of this is that VEGF-X can be used to repair the damage of myocardial and other occasions where coronary bypass surgery is needed by stimulating the growth of the transplanted tissue. Related to this is the use of VEGFX to repair the cardiac vascular system after ischemia.

The protein of the present invention may also be employed in accordance with the present invention by expression of such protein *in vivo*, which is often referred to as "gene therapy".

Thus, for example, cells such as bone marrow cells may be engineered with a polynucleotide (DNA or RNA) encoding for the protein *ex vivo* as defined herein, the engineered cells are then provided to a patient to be treated with the polypeptide. Such methods are well-known in the art. For example, cells may be engineered by procedures known in the art by use of a retroviral particle containing RNA encoding for the protein of the present invention.

Similarly, cells may be engineered *in vivo* for expression of the protein *in vivo*, for example, by procedures known in the art.

A further aspect of the invention comprises a method of treating a disorder mediated by expression of a protein according to the invention, by administering to a patient an amount of an antisense molecule as described herein, in sufficient concentration to alleviate or reduce the symptoms of said disorder.

Compounds which inhibit or enhance angiogenesis may be identified by providing a host cell or organism according to the invention or a transgenic cell, tissue or organism according to the invention, contacting a test compound with said cell, tissue or organism and monitoring for the effect of said compound compared to a cell tissue or organism which has not been contacted with said compound. These compounds may themselves be used as a medicament or included in a pharmaceutical composition for treatment of disorders mediated by inappropriate vascularisation or angiogenic activity.

The present inventors have also, advantageously, identified in the sequence encoding the VEGF-X protein a CUB domain, which has heretofore not previously been identified in VEGF-type growth factors. The VEGF-X protein may therefore exert dual regulatory effects via interaction with the VEGF tyrosine kinase receptors or with neuropilin receptors mediated by the CUB domain. Thus, the sequence encoding said CUB domain may be included in an expression vector for subsequent transformation of a host cell, tissue or organism.

VEGF-X or fragments thereof may be able to modulate

the effects of pro-angiogenic growth factors such as VEGF as indicated in the findings presented in the examples below that the N-terminal part of the VEGF-X protein, a CUB-like domain, is able to inhibit VEGF-stimulated proliferation of HUVECs. VEGF-X or fragments thereof may therefore be useful in therapy of conditions involving inappropriate angiogenesis. Inhibition of the angiogenic activity of VEGF has been linked with inhibition of tumour growth in several models eg Kim K. J. et al, Nature 362:841-844, (1993). Additionally, agents able to inhibit angiogenesis would be expected to be useful in treating other angiogenesis-dependent diseases such a retinopathy, osteoarthritis and psoriasis(Folkman, J., Nature Medicine 1:27-31, (1995).

As identified in more detail in the Examples described herein the present inventors have surprisingly identified that the CUB domain of VEGF-X is able to inhibit stimulation of proliferation of HUVECs induced by either VEGF or bFGF. The CUB domain may, therefore, be utilised as a therapuetic agent for inhibition of angiogenesis and for treatment of condition associated with inappropriate vascularisation or angiogenesis.

Therefore according to a further aspect of the invention there is provided a method of inhibiting angiogenic activity and inappropriate vascularisation including formation and proliferation of new blood vessels, growth and development of tissues, tissue regeneration and organ and tissue repair in a subject said method comprising administering to said subject an amount of a polypeptide having an amino acid sequence from position 40 to 150 of the sequence illustrated in Figure 10 or a nucleic acid molecule encoding the CUB domain according to the invention in

sufficient concentration to reduce or prevent said angiogenic activity.

Furthermore there is also provided a method of treating or preventing any of cancer, rheumatoid arthritis, psoriasis and diabetic retinopathy, said method comprising administering to said subject an amount of a polypeptide having an amino acid sequence from position 40 to 150 of the sequence illustrated in Figure 10 or a nucleic acid molecule encoding the CUB domain according to the invention in sufficient concentration to treat or prevent said disorders.

The CUB domain may also be used to identify compounds that inhibit or enhance angiogenic activity such as inappropriate vascularisation, in a method comprising contacting a cell expressing a VEGF receptor and/or a neuropilin 1 or 2 type receptor with said compound in the presence of a VEGF-X protein according to the invention and monitoring for the effect of said compound or said cell when compared to a cell which has not been contacted with said compound. Such compounds may then be used as appropriate to prevent or inhibit angiogenic activity to treat the disorders or conditions described herein, or in a pharmaceutical composition. An antibody to said CUB domain may also be useful in identifying other proteins having said sequences.

Deposited Plasmids

	Date of Deposit	Accession No.
Plasmid VEGFX/pCR2.1 1TOPO FL	1 March 1999	LMBP 3925
Plasmid VEGFX/pRSETB BD amino acids G230-G345	1 March 1999	LMBP 3926
Plasmid VEGFX/pcR.2.1 FL Clone 9	20 October 1999	LMBP 3977
Plasmid VEGF-X CUB PET22b	20 December 1999	------

The above plasmids were deposited at the Belgian Coordinated Collections of Microorganisms (BCCM) at Laboratorium Voor Moleculaire Biologie-Plasmidencollectie (LMBP) B-9000, Ghent, Belgium, in accordance with the provisions of the Budapest Treaty of 28 April 1977.

The invention may be more clearly understood with reference to the accompanying example, which is purely exemplary, with reference to the accompanying drawings, wherein:

Figure 1: is a DNA sequence identified in the Incyte LifeSeq™ database coding for a novel VEGF-X protein.

Figure 2: is an illustration of amino acid sequence of the nucleic acid sequence of Figure 1.

Figure 3: is an illustration of PCR primer sequences utilised to identify the VEGF-X protein according to the invention.

Figure 4: is a diagrammatic illustration of the spatial relationships in the VEGF-X sequence of the clones identified using the PCR primer sequences of Figure 3.

Figure 5: is an illustration of the nucleotide sequences of the 5′ RACE primers used to identify the 5′ end of the VEGF-X open reading frame.

Figure 6: is an illustration of the sequence obtained from the RACE experiment.

Figure 7: is an illustration of the nucleotide sequences obtained from the search of LifeSeq™ database using the sequence in Figure 6.

Figure 8: is an illustration of the primers used to clone the entire coding sequence of VEGF-X.

Figure 9: is an illustration of the entire coding sequence of VEGF-X.

Figure 10: is an illustration of the predicted amino acid sequence of the nucleotide sequence of Figure 9.

Figure 11: is an alignment of the sequence of

Figure 10 with the sequences of VEGF-A to D.

Figure 12: is an illustration of variant sequences of the VEGF-X protein according to the invention.

Figure 13: is an illustration of the oligonucleotide primers used for E.coli expression of VEGF-X domains and for expression of the full length sequence of VEGF-X in a baculovirus/insect cell expression system.

Figure 14: depicts nucleic acid sequences of 18 human EST clones obtained from a BLAST search of the LifeSeq™ database used to identify the full sequence encoding VEGF-X.

Figure 15: depicts the nucleotide sequences of 50 human EST clones obtained from the LifeSeq™ database.

Figure 16: is an illustration of nucleotide sequences utilised as primers to identify the nucleotide sequence encoding VEGF-X.

Figure 17: is a nucleotide sequence coding for a partial VEGF-X protein according to the invention.

Figure 18: is an illustration of a partial nucleotide sequence encoding VEGF-X protein according to the invention.

Figure 19: is an illustration of a DNA and polypeptide sequence used for mammalian cell expression of VEGF-X. The predicted VEGF-X signal sequence is in lower case letters. The C-terminal V5 epitope and His6 sequences are underlined.

Figure 20: is an illustration of a DNA and polypeptide sequence used for baculovirus/insect cell expression of VEGF-X. In the polypeptide sequence the signal sequence is shown in lower case. The N-terminal peptide tag added to the predicted mature VEGF-X sequence is underlined.

Figure 21: is an illustration of a DNA and polypeptide sequence used for *E. coli* expression of VEGF-X. The polypeptide sequences at the N- and C- termini derived from the MBP fusion and His6 tag respectively are underlined.

Figure 22: illustrates the disulphide-linked dimerisation of VEGF-X. Protein samples were analysed by SDS-PAGE. Prior to loading the gel, samples were heated to 95°C for 5 minutes in sample buffer in the presence (+) or absence (-) of reducing agent. (A) samples from COS cell expression of a C-terminally V5/His6 peptide-tagged construct. The left hand panel is total conditioned medium, the right hand panel is material purified on Nickel agarose resin. Reduced monomer

and putative disulphide-linked, non-reduced dimer are indicated by arrows. There appears to be proteolysis of the protein during purification. Gels were blotted onto nylon membranes and protein detected with an anti V5 monoclonal antibody. (B) Samples from *E.coli* expression of a maltose-binding protein/His6 dual fusion construct. M indicates the molecular weight markers (Benchmark, LifeTechnologies). The gel was stained with Coomassie Blue by standard procedures. The fusion protein has an apparent molecular weight of 80kDa.

Figure 23: illustrates the glycosylation of VEGF-X. VEGF-X was purified from the culture supernatant of COS cells transfected with the pcDNA6/V5-His construct. Supernatants were harvested 72h post-transfection and purified on nickel resin. Samples were then treated with EndoH (+) or untreated (-) before SDS-PAGE and blotting, as described in the legend to Figure 22.

Figure 24: is an illustration of the DNA and polypeptide sequence used for *E. coli* expression of the VEGF-like domain of VEGF-X. Polypeptide sequences at the N-terminus of the protein derived from the vector are underlined.

Figure 25: shows expression of the VEGF-X VEGF domain in *E. coli*. Lane 1-10µl broad

range marker (New England Biolabs), lane 2-10µl unreduced sample, lane 3-10µl reduced sample. The reduced PDGF domain protein (lane 3) has an apparent molecular weight of approximately 19kDa on SDS-PAGE.

Figure 26: illustrates a DNA and polypeptide sequence used for *E. coli* expression of the CUB-like domain of VEGF-X. The polypeptide sequence at the N-terminus derived from the vector-encoded signal and the introduced His6 tag are underlined.

Figure 27: shows expression of the VEGF-X CUB domain in *E. coli*. The CUB domain protein was purified on Nickel chelate resin. The protein migrates at approximately 23kDa on SDS-PAGE.

Figure 28: illustrates the effect of truncated VEGF-X (CUB domain) on HUVEC proliferation. (A) Human Umbilical Vein Endothelial Cells (one-day-treatment). (B) Human Umbilical Vein Endothelial Cells (24-hour starving followed by one-day-treatment). (C) Effect of VEGF-A_{165} and VEGF-X CUB domain on the proliferation of HUVEC (two-day-treatment).

Figure 29: depicts the tissue distribution of VEGF-X mRNA analysed by Northern blotting and RT-PCR in (A) normal tissues and (B) tumour tissue and cell lines.

Figure 30: depicts the partial intron/exon structure of the VEGF-X gene. (A) Genomic DNA sequences of 2 exons determined by sequencing; exon sequence is in upper case, intron sequence is in lower case. (B) Shows the location of splice sites within the VEGF-X cDNA sequence. The location of mRNA splicing events is indicated by vertical lines. The cryptic splice donor/acceptor site at nt. 998/999 (diagonal lines) gives rise to the splice variant forms of VEGF-X. No splice site information is given for the region shown in italics.

Figure 31: is a graphic representation of the effect of FL-VEGF-X on HuVEC proliferation: (24 hour serum starvation followed by one day treatment).

Figure 32: is a graphic representation of the combined effect of truncated VEGF-X (CUB domain) and human recombinant $VEGF_{165}$ on HuVEC proliferation: (24 hour serum starvation followed by two day treatment).

Figure 33: is a graphic representation of the combined effect of the CUB domain and human recombinant bFGF on HuVEC proliferation: (24 hour serum starvation followed by two day treatment).

Figure 34: is a graphic representation of the

results of a LDH assay for testing cytotoxicity of the CUB domain or the CUB domain with rhVEGF$_{165}$.

Figure 35: is a graphic representation of the results obtained from a LDH assay for testing cytotoxicity of the CUB domain or CUB domain with rh-bFGF.

A BLAST (Basic Local Alignment Search Tool; Altschul *et al.*, 1990 J. Mol. Biol. 215, 403-410) search was performed in the proprietary LifeSeq™ human EST database (Incyte Pharmaceuticals, Inc., Palo Alto, CA, USA). BLAST produces alignments of both nucleotide and amino acid sequences to determine sequence similarity. Because of the local nature of the alignments, BLAST is especially useful in determining exact matches or in identifying homologues. While it is useful for matches which do not contain gaps, it is inappropriate for performing motif-style searching. The fundamental unit of BLAST algorithm output is the High-scoring Segment Pair (HSP).

Eighteen human EST clones (Figure 14) with high similarity to the previously identified VEGF proteins were identified and a further fifty EST clones (Figure 15) were identified using these sequences as query sequences, allowing us to deduce the putative sequence for the new VEGF-X protein. The sequences obtained were compared to known sequences to determine regions of homology and to identify the sequence as a novel VEGF-type protein. Using the DNA sequence information in the databases we were able to prepare suitable primers having the sequences of VEGF-X 1-10 illustrated in Figure 3 for use in subsequent RACE experiments to obtain the complete

DNA sequence for the VEGF-X gene.

Cloning

A profile was developed based on the VEGF-like domain in existing VEGF sequences (VEGF-A, B, C and D). This was used to search the public databases and the Incyte LifeSeq™ database. No significant novel matching sequences were found in the public databases. All of the matching sequences found in the LifeSeq™ database (~1000) were assembled to give a smaller number of sequences (~30), which included the known VEGFs and a potential novel VEGF (**figures 1 and 2**). This sequence was named VEGF-X.

Oligonucleotides were designed to amplify the VEGF-X sequence from cDNA (**figure 3**). The ESTs found in LifeSeq™ were from a range of tissues, with a slight predominance of sequences from ovary, testis, placenta and lung (Figure 14 and 15). Accordingly the oligonucleotides were used to amplify cDNA derived from lung and placenta. First-round PCR products were found at ~200bp larger than the expected sizes, while 3 major species appeared after a second round of PCR amplification, the smallest of which was of the expected size. These fragments were cloned and sequenced. The smallest fragment did indeed have the sequence originally identified from the LifeSeq database, while the others contained insertions (**figure 4**).

As the first round of amplification suggested that the major species found in cDNA from ovary and placenta was not that originally identified in the LifeSeq™ database, the focus of effort was switched to the presumed major species (it seemed likely that

clones 57, 25-27 and 2.1kb clones 1-3 in fig 4 represented the major mRNA species). Conceptual translation of the DNA sequences of these cloned PCR fragments indicated that the complete open reading frame was not present in the clones or in the sequence from LifeSeq™. While all clones contained the same sequence in the region of the translation termination codon, indicating that the end of the open reading frame had been identified, the 5′ end of the open reading frame had not been cloned. 5′ RACE experiments were therefore carried out in order to find the start of the reading frame. PCR primers designed for RACE experiments are shown in **figure 5**. RACE PCR products were sequenced directly. Sequence could be obtained from the 3′ end of these RACE products but not from the 5′ end; probably because the products were not cloned and were therefore heterogeneous at the 5′ end. This new sequence was assembled with the existing cloned sequence to give the sequence shown in **figure 6**. Searching the LifeSeq™ database with this sequence identifies ESTs which extend the sequence a further 140bp in the 5′ direction and a further 160bp in the 3′ direction (**figure 7**). This longer contig was used to design oligonucleotide primers to amplify the entire coding sequence (these primer sequences are shown in **figure 8**). PCR was carried out using primers 5′-1 and vegfX10 (in order to clone a "full-length" cDNA), and with primers 5′-1 and vegfX6 (in order to clone the full coding region, see **figure 3** for sequences of vegfX10 and vegfX6). A number of clones were obtained for the shorter fragment, of which clones 4 and 7 contain no PCR errors (sequence of clones 4 & 7 in **figure 9**). A single clone was obtained for the longer fragment (clone 9), but this sequence appears to contain 2 PCR errors.

The predicted polypeptide from these longer contigs is shown in **figure 10**. Amino acids 1-22 are predicted to encode a signal sequence (von Heijne, 1986, *Nucleic Acids Res.* **14**, 4683-4690). **Figure 11** shows an alignment of the protein sequence with VEGFs A-D. The region homologous to the other VEGFs is located towards the C-terminus of the protein. As the VEGF homology domain is expected to belong to the TGF-beta superfamily of growth factors and to consist of a dimer containing both intra- and intermolecular disulphide bonds, initial alignments focussed on the cysteines. However, mapping of the sequence onto the known x-ray structure of the VEGF-A receptor-binding domain (Muller et al (1997) *Proc. Natl. Acad. Sci USA* **94**, 7192-7197) suggests that the alignment in **figure 11** is plausible, as the extra 4 cysteine residues within the VEGF-homology region of VEGF-X (compared to this region of VEGF-A) correspond to residues which are spatially close in VEGF-A, and may therefore be able to form disulphide bonds.

A search of the PFAM database of protein domains with the full-length polypeptide sequence from **figure 10** identifies two domain consensus sequences within the polypeptide. The more C-terminal domain is a "VEGF" domain: (the known VEGFs all contain this domain and the structure of this region of VEGF-A is similar to that of PDGF). Additionally towards the N-terminus of the polypeptide there is a CUB domain (amino acids ~40-150). The CUB domain is a 100-110 amino acid extracellular domain found in a number of developmentally-regulated proteins. When the full-length protein is used to search the protein databases using the BLAST 2 algorithm, the scores for matches to CUB domain-containing proteins are more

significant than those to the other VEGFs. Interestingly, the most significant matches are to the CUB domains of Neuropilins, and Neuropilin-1 was recently identified as a receptor of one of the VEGF-A isoforms VEGF-A_{165} (Soker et al. (1998) *Cell* **92**, 735-745).

Assuming that the variant sequences isolated by PCR (i.e. the smaller PCR fragments) use the same translation initiation site as the full-length sequence, they would result in production of the variant proteins shown in **figure 12**. It may be significant that both of these variant proteins retain the CUB domain and delete all or part of the VEGF-like domain. The production of these variant sequences can be explained by the use of a cryptic splice donor/acceptor site within the VEGF-X sequence (figure 30B, between nt. 998/999): one variant arises by splicing out of the region between nt. 729-998, the other by splicing out of the region between nt. 999-1187.

Expression

Full-length expression constructs

Mammalian cells

Clone 4 containing the full CDS of VEGF-X (see figure 9), was used to generate constructs for expression of full-length protein. The sequence was amplified by PCR and cloned into the vector pCDNA6/V5-His so as to add a C-terminal V5 epitope tag and His_6 tag. The DNA and polypeptide sequence in this vector is shown in figure 19. Transient expression in COS cells followed by western blotting and detection via an anti-V5 mAb demonstrates the secretion of a protein of ~50K into the medium in transfected cells only

(figure 22A). This construct can also be used to generate VEGF-X expressing stable CHO cell lines.

Baculovirus/Insect-cell expression system

For expression in the baculovirus/insect cell system the DNA encoding the predicted mature VEGF-X polypeptide sequence was fused to a sequence encoding a signal derived from melittin, a secreted insect protein. An N-terminal 6His tag was also added to facilitate purification. The insert was then cloned into the baculovirus expression vector pFASTBAC. The DNA and polypeptide sequence of this construct is shown in figure 20. Infection of *Trichoplusia ni* Hi5 cells with this recombinant baculovirus results in the secretion of a protein of approximately 45K into the medium (data not shown).

E.coli

The coding region of VEGF-X has been cloned in a variety of ways for expression as a secreted protein in *E.coli*. A particularly useful expression clone carries an N-terminal fusion to the *E.coli* maltose-binding protein (MBP- derived from the expression vector pMAL-p2, New England Biolabs) and a C-terminal fusion to a 6His tag. The DNA and polypeptide sequence of this vector is shown in figure 21. Sequential purification of cell fractions on Ni-NTA resin and amylose resin allows the isolation of the expressed protein (see figure 22B).

Expression of fragments

VEGF

The VEGF domain of VEGF-X has been expressed in *E.coli*. Similar domains from VEGF-A (Christinger et al. (1996) *PROTEINS: Structure, Function and Genetics* 26, 353-357), and VEGF-D (Achen et al (1998) *Proc.*

Natl. Acad. Sci USA **95**, 548-553) have been shown to be capable of binding to the respective receptors. Expression of these domains was carried out using the bacterium *E.coli*. Additionally, the full-length protein was expressed using the baculovirus/insect cell expression system. The oligonucleotide primers which have been obtained for these experiments are shown in **figure 13**. The construct directed expression in the bacterial cytoplasm, and as expected the protein was produced in insoluble form in inclusion bodies (the DNA and polypeptide sequence used for PDGF domain expression is shown in figure 24). Inclusion bodies were washed, solubilized with urea and the protein purified under denaturing conditions, before refolding by dialysis to remove the urea. Soluble protein was obtained, but shows little evidence of the disulphide bond linked dimers seen with material derived from animal cells (figure 25, compare with figure 22A & B). It is not clear therefore whether this protein is correctly folded.

CUB

The CUB domain has been expressed as a soluble secreted protein in *E.coli* (figure 26). The protein was purified by binding to Ni-NTA resin (figure 27) and assayed for activity on HUVECs in an in-vitro proliferation assay.

Properties of the VEGF-X protein

The transient mammalian cell expression system described above has been used to generate full-length VEGF-X protein, as shown by antibody detection following Western blotting (see figure 22A).

Disulphide bond linked dimers

The other members of the PDGF family of growth

factors, the PDGFs and VEGFs, all exist as dimers in which two monomers constituting the dimer are linked by interchain disulphide bonds. The x-ray structures of PDGF-BB (Oefner et al, 1992), and VEGF-A (Muller et al, 1997) are known and indicate that at least these two members of the family contain two interchain disulphide bonds. Practically this means that in SDS-PAGE analysis of these growth factors the presence of interchain disulphide bonds is shown by a large decrease in mobility in the absence of reducing agent (ie. the nonreduced dimer migrates more slowly through the gel than the reduced monomer). This effect was also expected for VEGF-X, and has been demonstrated for the material obtained from transient mammalian cell expression (figure 22A). In the case of the full length material produced in *E.coli* only some 10% of the total VEGF-X protein appears to be present as disulphide bond-linked dimers (figure 22B). However, these results provide evidence that the mammalian cell-derived protein is correctly folded, and that a portion of the *E.coli*-derived protein is too.

Glycosylation

There are 3 predicted potential N-linked glycosylation sites within the VEGF-X protein: at residues 25, 55 and 254 of the polypeptide sequence. The predicted molecular mass of the mature VEGF-X protein is 40kDa, but SDS-PAGE and western blotting (detection via an introduced C-terminal epitope tag- see figure 19) of the full-length protein expressed in COS cells gives a band slightly larger than the expected size (45-50kDa) as well as one at 25kDa (figure 22A). This smaller band is presumed to be a C-terminal proteolysis fragment derived from the full-length molecule (controls from uninfected cells do not show this band), probably corresponding to a

cleavage between the CUB and VEGF domains. EndoH treatment of the preparation gives a slight mobility change for the full-length protein (figure 23), but for the smaller VEGF domain fragment there is a clear change, indicating that the predicted glycosylation site within the VEGF domain at residue 254 is indeed glycosylated.

Activity of proteins in cell-based assays

Protein samples were tested for activity in cell proliferation, cell migration and *in-vitro* angiogenesis assays. Active samples can also be tested in the *in vivo* matrigel mouse model of angiogenesis.

Full-length VEGF-X protein

Conditioned medium derived from COS cells transiently expressing VEGF-X (see figure 22A) displayed no detectable activity in any of the assays. However, as VEGF-X protein could only be detected in this preparation by Western blotting, and not by Coomassie-staining of gels, it is clearly present at very low levels and this may be the reason for the observed lack of activity in the cell proliferation, migration or *in vitro* angiogenesis tests.

VEGF domain

The VEGF domain protein described above has been tested in cell proliferation (on a range of cell types), cell migration and in vitro angiogenesis assays and has failed to show activity in any of these tests. As suggested above, this may be due to incorrect folding of this protein.

CUB domain

The CUB domain protein at the highest dose tested

(1µg/ml) appears to inhibit proliferation of HUVECs in the absence of other stimulation (figure 28A & B). This effect is also seen following stimulation with the lowest VEGF-A_{165} dose tested (1ng/ml- figure 28C). The CUB domain of VEGF-X therefore appears to show antiproliferative activity on HUVECs, even in the presence of low VEGF-A_{165} doses.

Tissue distribution of mRNA

VEGF-A mRNA expression has been shown to be upregulated in a wide variety of human tumors (lung, breast, ovarian, colon, stomach, liver, pancreas, kidney, bladder and prostate- Takahashi et al, 1995). Tumor VEGF-A expression has been shown to correlate with tumor growth rate, microvascular density and tumor metastasis (Takahashi et al, 1995). It was thus of interest to examine the mRNA expression patterns of VEGF-X. Accordingly, Northern blot analysis of mRNA derived from different tissues has been carried out. The results indicate that although the VEGF-X mRNA is expressed at low levels, it is present in a wide range of tissues. PCR amplification of cDNA from a range of tissue sources supports this idea (figure 29A). The major mRNA species is approximately 3.1kb in size. There is no significant upregulation seen in tumour cell lines or in tumour tissues tested (figure 29B), with the possible exception of the cell lines GI-117 (lung carcinoma) and SaOS-2 (osteosarcoma). The results of these initial tissue distribution studies do not, therefore, provide evidence for upregulation of VEGF-X in tumour growth, as is seen with VEGF-A.

Genomic structure of the VEGF-X gene

A genomic BAC clone covering the 3′ part of the VEGF-X locus was isolated by hybridisation screening

of nylon filters containing a human BAC library. Direct sequencing of this clone using oligonucleotide primers based on the VEGF-X cDNA sequence allowed the determination of several intron/exon boundaries (figure 30). Interestingly, the position of the mRNA splice site within the PDGF domain (nt 1187/1188 in figure 30B) is conserved with respect to those in the VEGF-A and VEGF-D genes (Tischer et al, 1991; Rocchigiani et al, 1998).

Materials & Methods

PCR, Cloning, DNA sequence determination and BAC screening.

All primers were purchased from Eurogentec, Seraing, Belgium. Insert-specific sequencing primers (15- and 16-mers) were designed by visual inspection of the DNA sequences. DNA was prepared on Qiagen-tip-20 columns or on Qiaquick spin columns (Qiagen GmbH, Düsseldorf, Germany) and recovered from the spin columns in 30µl Tris/EDTA-buffer (10mM TrisHCl pH 7.5, 1 mM EDTA (sodium salt)). Sequencing reactions were performed using BigDye™ Terminator Cycle Sequencing Ready Reaction kits (Perkin Elmer, ABI Division, Foster City, CA, USA) and were run on an Applied Biosystems 377 DNA sequencer (Perkin Elmer, ABI Division, Foster City, CA, USA). Polymerase chain reactions were carried out according to standard procedures (Ausubel et al, 1997). The PCR fragments were cloned into vectors pCR2.1 (Invitrogen, Carlsbad, CA. USA) or pCR-TOPO (Invitrogen,NL) according to the manufacturer's instructions. One of those vectors, plasmid VEGFX/pCR2.1 1TOPO FL was deposited on 1 March 1999 under Accession No. LMBP 3925. After sequence determination, the inserts were cloned into the desired expression vectors (see

figures 19, 20, 21, 24 & 26).

A human genomic BAC library (Genome Systems, Inc., St Louis, MI, USA) was screened by hybridisation to oligonucleotides derived from the VEGF-X cDNA sequence, according to the manufacturer's instructions. BAC DNA was prepared using a Qiagen plasmid midi kit (Qiagen GmbH, Düsseldorf, Germany) according to the manufacturer's instructions with some modifications (after clearing of the lysate from chromosomal DNA, supernatants from individual preparations were pooled on a single column (tip 100), and after the 70 % EtOH wash, the pellet was resuspended overnight at 4°C in 100 µl TE). 20-mer sequencing primers were designed based on the known cDNA sequence, and sequencing carried out as above.

5' RACE

In order to extend the cDNA clone in a 5' direction RACE reactions were carried out. Since it was known that the mRNA is present in placenta and skeletal muscle, Marathon-Ready™ placenta and skeletal muscle cDNAs were purchased from Clontech (Palo Alto CA. USA) and used according to the manufacturer's instructions. DNA fragments were excised from agarose gels, purified using QiaQuick PCR purification columns (Qiagen GmbH, Düsseldorf, Germany) and sequenced directly.

VEGF-X protein expression and purification

DNA fragments encoding the desired protein sequences were amplified by PCR and cloned into appropriate expression vector systems.

For mammalian cell expression, the full coding

sequence was cloned into the vector pcDNA6/V5-his (Invitrogen Leek, NL, see figure 19 for construct sequence), so as to add a C-terminal peptide tag to assist in detection and purification.

For insect cell expression the sequence of the predicted mature polypeptide was initially amplified to add an N-terminal 6His peptide and then cloned into the pMelBacB vector (Invitrogen, Leek, NL) to add an insect cell signal sequence. The entire insert was then PCR-cloned into the vector pFASTBAC-1 (LifeTechnologies, Gaithersburg, MA, USA) for construction of a baculovirus according to the manufacturer's instructions.

For *E.coli* expression, the coding region was PCR amplified to add a C-terminal 6His tag and then cloned into the vector pMAL-p2 (New England Biolabs, Beverly, MA, USA). The coding sequence of this construct is shown in figure 21). The protein was purified first on Ni-NTA resin (Qiagen GmbH, Düsseldorf, Germany) and then on amylose resin (New England Biolabs, Beverly, MA, USA), according to the manufacturers' instructions.

DNA sequences encoding the CUB and VEGF domain fragments of VEGF-X were PCR amplified and cloned into pET22b and pET21a (Novagen, Madison, WI, USA) respectively. The CUB domain protein was prepared either from the periplasm or medium of induced cultures by standard methods (Ausubel et al, 1997). The protein was initially purified by precipitation with 20% ammonium sulphate. After overnight dialysis vs 20mM Tris Hcl pH7.5, 100mM NaCl to remove ammonium sulphate, the protein was further purified on Ni-NTA resin as described above. The VEGF domain protein was expressed in insoluble form, and preparation of

inclusion bodies was carried out using standard procedures (Ausubel et al 1997). Inclusion bodies were dissolved in 6M guanidine hydrochloride, 20mM Tris Hcl pH8.0, 200mM NaCl, 1mM 2-mercaptoethanol, and purified on Ni-NTA resin (Qiagen GmbH, Düsseldorf, Germany) according to the manufacturer's instructions. The protein was refolded by dialysis against several changes of buffer containing decreasing concentrations of denaturant.

Analysis of protein glycosylation was carried out using EndoH (Roche Molecular Biochemicals, Brussels, BE) according to the manufacturer's instructions.

Cell Proliferation Assay

Human umbilical vein endothelial cells (HUVECs) (Clonetics, San Diego, CA.) were trypsinized with 0.05% trypsin/0.53mM EDTA (Gibco, Gaithersburg, MD.), resuspended in the EGM-2(Clonetics, San Diego, CA.), counted, and distributed in a 96-well tissue culture plate at 5,000 cells/well. Following cell attachment and monolayer formation (16 hours), cells were stimulated with various concentrations of truncated VEGF-X (CUB domain or VEGF domain) or dilutions of culture supernatants of the full-length VEGF-X (COS 7 or HEK293) in DMEM (Gibco, Gaithersburg, MD.) containing 0.5% to 2% FBS (HyClone, Logan, UT) as indicated. For human fetal dermal fibroblasts (American Type Culture Collection, Rockville, MD.), the growth medium was replaced by DMEM containing 0.1% BSA (Sigma, St. Louise, MO.) with or without various concentrations of truncated VEGF-X proteins. For HCASMC (Clonetics, San Diego, CA.), the medium was replaced by DMEM containing 0.5% FBS. The cells were treated for a further 24 hr-72 hr. For the measurement of proliferation, the culture media were replaced with 100 µl of DMEM containing 5% FBS and 3

μCi/ml of [3H]-thymidine (Amersham, Arlington Heights, IL.). Following pulse labeling, cells were fixed with methanol/acetic acid (3:1, vol/vol) for 1 hour at room temperature. The cells were washed twice with 250 μl/well of 80% methanol. The cells were solubilized in 0.05% trypsin (100μl/well) for 30 minutes then in 0.5% SDS (100 μl/well) for another 30 minutes. Aliquots of cell lysates (180 μl) were combined with 2 ml of scintillation cocktail (Fisher, Springfiled, NJ) and the radioactivity of cell lysates was measured using a liquid scintillation counter (Wallac 1409). In each case, samples were performed in quadruplicate.

Chemotaxis Assay

The chemotactic response of HUVECs was assayed using a 48-well modified Boyden chamber (NeuroProbe, Cabin John, MD.) and collagen-coated (0.1mg/ml type I collagen, Collaboratic Biomedical, Bedford, MA.) polycarbonate membrane filters with a pore diameter of 8 μm (NeuroProbe, Cabin John, MD.). Cell suspensions (15,000/well) were loaded to the upper part of the chemotaxis chamber and stimulated for 4 hours with rhVEGF$_{165}$ (0.1-10 ng/ml) (Calbiochem, San Diego, CA.) or various concentrations of truncated VEGF-X (PDGF domain). Cells remaining on the top of the membrane were removed. Migration was assessed by counting the number of cells that migrated to the lower side of the filter membrane. The membrane was fixed with 10% formaldehyde for 15 min, followed by staining with Gill's hemotoxylin III (Poly Scientific, Bay Shore, NY.). The assay was performed in triplicates and six independent high power fields per well were counted using a light microscope at 250 magnification. The results were expressed as the fold of unstimulated cells (EGM containing 0.1% BSA).

***In Vitro* Angiogenesis Assay**

In vitro angiogenesis in fibrin gels was quantitated using spheroids of human umbilical vein endothelial cells (Korff et al., 1998). To generate endothelial cell spheroids of defined size and cell number, a specific number of cells (~ 800 cells per spheroid) was suspended in EGM-2 culture medium containing 20% methylcellulose (Sigma, St. Louis, MO.), seeded into nonadherent round-bottom 96-well plates. All suspended cells in one well contributed to the formation of a single endothelial cell spheroid within 24 hours. A fibrin gel stock solution was prepared freshly prior to use by mixing 3mg/ml fibrinogen (Calbiochem, San Diego, CA.) in Medium 199(Gibco, Gaithersburg, MD.). Assays were performed in 24-well culture plates. The 1ml fibrinogen stock was mixed with 50 HUVEC spheroids and the corresponding test substance including rh-$VEGF_{165}$ or various concentration of VEGF-X. The spheriod-containing fibrinogen was rapidly transferred into 24-well plates. Fifteen microliters of thrombin (100 NIH U/ml stock, Sigma, St. Louis, MO.) was added to the gel for the fibrin gel formation. The gel formation usually occurred within 30 seconds. After gel formation, 1ml/well of Medium 199 supplemented with 20% FBS, 1mg/ml ε-aminocaproic acid (Calbiochem, San Diego, CA.) and antibiotics were added. The gel was incubated at 37°C (5%CO_2, 95% air, 100% humidity). After 3 days, *in vitro* angiogenesis was quantitated by measuring the length of the three longest capillary sprouts that had grown out of each spheroid (100X magnification), analyzing at least 10 spheroids per experimental group and experiment.

Matrigel Mouse Assay

The matrigel mouse assay is carried out as described by Passanti et al (1992).

Analysis of VEGF-X gene expression by RT-PCR analysis.

Oligonucleotide primers VEGF-E2 and VEGF-X14 (figure 16; figure 5) were used for the specific PCR amplification of a 350 bp fragment from VEGF-X. PCR amplifications were performed on human multiple tissue cDNA (MTC™) panels (Clontech human MTC panels I and II and human Tumor MTC panel) normalised to the mRNA expression levels of six different housekeeping genes. In addition, cDNA was made from different tumor cell cultures (Caco-2 colorectal adenocarcinoma; T-84 colorectal carcinoma; MCF-7 breast adenocarcinoma; T-47D breast ductal gland carcinoma; HT1080 bone fibrosarcoma; SaOS-2 osteosarcoma; SK-N-MC neuroblastoma; HepG2 hepatoblastoma; JURKAT T-cell leukemia and THP-1 myelomonocytic leukemia). For the preparation of tumor cell cDNA, cells were homogenised and total RNA prepared using the RNeasy Mini kit (Qiagen GmbH, Hilden, Germany) according to manufacturer's instructions. 1 µg of total RNA was reverse transcribed using oligo(dT)15 as a primer and 50 U of Expand™ Reverse Transcriptase (Boehringer Mannheim, Mannheim, Germany) according to the manufacturer's instructions. PCR reactions with VEGF-X-specific or glyceraldehyde-3-phosphate dehydrogenase (G3PDH)-specific primers were then performed on 1 µl of this cDNA. For all cDNAs, PCR reactions with VEGF-X specific primers were performed in a total volume of 50 µl, containing 5 µl (± 1 ng) of cDNA, 1x Advantage KlenTaq PCR reaction buffer, 0.2 mM dNTP, 250 nM of primers VEGF-E2 and VEGF-X14 and 1 µl of Advantage KlenTaq polymerase mix. Samples were heated

to 95°C for 30 s and cycling was done for 30 s at 95°C and 30 s at 68°C for 25, 30 or 35 cycles. Control reactions using specific primers that amplify a 1 kb fragment of the housekeeping gene G3PDH were also performed according to the manufacturer's instructions.

Northern blot analysis of VEGF-X.

Northern blots containing 2 µg of poly(A)-rich RNA derived from different human tissues (Clontech Laboratories; MTN™ blot, MTN™ blot II and Cancer Cell Line MTN™ blot) were hybridized according to the manufacturers instructions with a α-[^{32}P]-dCTP random-priming labelled (Multiprime labelling kit, Roche Diagnostics) 293 bp specific VEGF-X fragment (*Pin*AI-*Stu*I fragment including 92 bp of the 3' end coding region and 201 bp of the 3' untranslated region of VEGF-X). The blots were hybridized overnight at 68°C and final washes at high stringency were at 68°C in 0.1x SSC/0.1 % SDS. The membranes were autoradiographed for 1 to 3 days with intensifying screens.

Full length VEGF-X

The effect of full length VEGF-X on proliferation of HuVEC cells was determined by the ^{3}H-Thymidine incorporation assay. HuVEC cells were serum starved for 24 hours prior to treatment with the full length VEGF-X at the concentration range from 100 pg/ml-10 µg/ml. There was no effect of VEGF-X at 100 pg/ml-10 ng/ml on endothelial cell proliferation. At the higher concentrations of FL-VEGF-X (100 ng/ml and 1 µg/ml) there was a marked inhibition of endothelial cell proliferation. This is probably due to the very high endotoxin level in the samples. The VEGF-X sample was purified in order to decrease the

endotoxin level and is currently tested in the cell proliferation assay.

The Summary from Testing the CUB Domain

The effect of CUB domain on inhibition of HuVEC prolieration either serum- (2%), rh-VEGF or bFGF-stimulated, was assessed by the ^{3}H-Thymidine incorporation assay. Cells were serum starved followed by the treatment with the CUB domain and various growth factors. Results showed that the CUB domain inhibited endothelial cell proliferation, either serum- (2%), rh-VEGF or bFGF-stimulated in a dose dependent manner with maximal inhibition at 10 µg/ml. There was approximately a 2-fold inhibition of proliferation (at 10 µg/ml) of cells stimulated with VEGF and bFGF and nearly a 5-fold inhibition of cells stimulated with serum (2%). Results with the LDH assay showed that there was no cytotoxicity associated with the inhibition of cell proliferation by the CUB domain.

Therefore, the N-terminus of the polypeptide from Figure 10 has been shown to possess a CUB domain. When database searches are carried out using the full-length coding sequence the best matches (i.e. for a BLAST search, those with the lowest probability score) are found with the CUB domain rather than with the VEGF-like domain. The best match from searching release 37 of the SWISSPROT database (Feb 1999) is to the CUB domain of a neuropilin from Xenopus laevis, and the matches to the CUB domains of human neuropilins 1 and 2 are also more significant than matches to the VEGFs.

This similarity is provocative, given the identification of neuropilin-1 and -2 as cellular receptors for the VEGF-A 165 (Stoker *et al.* 1998,

reviewed in Neufeld *et al.* 1999). It is plausible therefore that VEGF-X could exert dual regulatory effects: via interaction with the tyrosine kinase VEGF-receptors mediated by the VEGF-like domain, as well as via interaction with VEGF isoforms or with the neurophilin receptors, mediated by the CUB domain.

To the best of our understanding the latter would be entirely novel, and searches on the most recent release of the Incyte database do not reveal any other proteins containing both CUB and VEGF-like domains. This arrangement of domains suggests possible positive or negative models of regulation:

Positive- the VEGF-like domain is able to interact productively with the tyrosine kinase VEGF receptors giving activation, and the CUB domain is able to interact productively with the neuropilin receptor giving activation.

Negative- the VEGF-like domain does not interact productively with the tyrosine kinase VEGF receptors, either preventing receptor dimerisation or blocking the VEGF binding sites. Further, the CUB domain does not interact productively with the neuropilin receptors, either preventing receptor activation or blocking the VEGF binding sites, or indeed by binding to VEGF isoforms and preventing their interaction with receptors.

TABLE 1

ORIGINAL RESIDUE	EXEMPLARY SUBSTITUTIONS
ALA	SER, THR
ARG	LYS
ASN	HIS, SER
ASP	GLU, ASN
CYS	SER
GLN	ASN, HIS
GLU	ASP, GLU
GLY	ALA, SER
HIS	ASN, GLN
ILE	LEU, VAL, THR
LEU	ILE, VAL
LYS	ARG, GLN, GLU, THR
MET	LEU, ILE, VAL
PHE	LEU, TYR
SER	THR, ALA, ASN
THR	SER, ALA
TRP	ARG, SER
TYR	PHE
VAL	ILE, LEU ALA
PRO	ALA

References

1. Ausubel, FM, R Brent, RE Kingston, DD Moore, JG Seidman, JA Smith, K Struhl (Eds). (1997) *Current Protocols in Molecular Biology*, John Wiley and Sons.

2. von Heijne, G. (1986) *Nucleic Acids Res.* **14**, 4683-4690.

3. Muller, YA, B Li, HW Christinger, JA Wells, BC Cunningham and AM de Vos. (1997) Vascular endothelial growth factor: crystal structure and functional mapping of the kinase domain receptor binding site. *Proc. Natl. Acad. Sci USA* **94**, 7192-7197.

4. Korff, T and Augustic, H.G. (1998) Integration of endothelial cells in multicellular spheroids prevents apoptosis and induced differentiation. *The Journal of Cell Biology.* 143, 1341-1352

5. Christinger, HW, YA Muller, LT Berleau, BA Keyt, BC Cunningham, N Ferrara and AM de Vos. (1996) *PROTEINS: Structure, Function and Genetics* **26**, 353-357.

6. Achen, MG, M Jeltsch, E Kukk, T Makinen, A Vitali, AF Wilks, K Alitalo and SA Stacker. (1998) *Proc. Natl. Acad. Sci USA* **95**, 548-553.

7. Siemeister, G, B Schnurr, K Mohrs, C Schachtele, C Marme and G Martiny-Baron. (1996) *Biochem. Biophys. Res. Commun.* **222**, 249-255.

8. Soker, S, S Takashima, HQ Miao, G Neufeld and M

Klagsbrun (1998). Neuropilin-1 is expressed by endothelial and tumor cells as an isoform-specific receptor for vascular endothelial growth factor, *Cell 92*: 735-745.

9. Neufeld, G, T Cohen, S Gengrinovitch and Z Poltorak (1999). Vascular endothelial growth factor and its receptors, *FASEB J.* 13:9-22.

10. Oefner, C., D'Arcy, A., Winkler, F.K., Eggimann, B. and Hosang, M. (1992). Crystal structure of human platelet-derived growth factor BB. *EMBO J.* 11, 3921-3926.

11. Passanti, A.., Taylor, R.M., Pili, R., Guo, Y., Long, P.V., Haney, J.A., Pauly, R., Grant, D.S. and Martin, G.R. (1992) A simple, quantitative method for assessing angiogenesis and antiangiogenic agents using reconstituted basement membrane, heparin and fibroblast growth factor. *Laboratory Investigation*, 67, 519-528.

12. Rocchigiani, M., Lestingi, M., Luddi, A., Orlandini, M., Franco, B., Rossi, E., Ballabio, A., Zuffardi, O. and Oliviero, S. (1990). Human FIGF: cloning, gene structure, and mapping to chromosome Xp22.1 between the PIGA and the GRPR genes. *Genomics*, 47, 207-216.

13. Takahashi, Y., Kitadai, Y., Bucana, C.D., Cleary, K.R. and Ellis, L.M. (1995). Expression of vascular endothelial growth factor and its receptor, KDR, correlates with vascularity, metastasis and proliferation of human colon

cancer. *Cancer Research*, 55: 3964-3968.

14. Tischer, E., Mitchell, R., Hartman, T., Silva, M., Gospodarowicz, D., Fiddes, J.C. and Abraham, J.A. (1991). The human gene for vascular endothelial growth factor: Multiple protein forms are encoded through alternative exon splicing. *J. Biol. Chem.* 266, 11947-11954.

SEQUENCE LISTING

Sequence ID No 1 corresponds to the amino acid sequence from position 23 to 345 of the amino acid sequence illustrated in Figure 10.

Sequence ID No 2 is the amino acid sequence illustrated in Figure 10.

Sequence ID No 3 corresponds to the sequence from position 257 to 1291 of the nucleotide sequence illustrated in Figure 9.

Sequence ID No 4 corresponds to the polynucleotide sequence of VEGFX1 illustrated in Figure 3.

Sequence ID No 5 corresponds to the polynucleotide sequence of VEGFX2 illustrated in Figure 3.

Sequence ID No 6 corresponds to the polynucleotide sequence of VEGFX3 illustrated in Figure 3.

Sequence ID No 7 corresponds to the polynucleotide sequence of VEGFX4 illustrated in Figure 3.

Sequence ID No 8 corresponds to the polynucleotide sequence of VEGFX5 illustrated in Figure 3.

Sequence ID No 9 corresponds to the polynucleotide sequence of VEGFX6 illustrated in

Figure 3.

Sequence ID No 10 corresponds to the polynucleotide sequence of VEGFX7 illustrated in Figure 3.

Sequence ID No 11 corresponds to the polynucleotide sequence of VEGFX8 illustrated in Figure 3.

Sequence ID No 12 corresponds to the polynucleotide sequence of VEGFX9 illustrated in Figure 3.

Sequence ID No 13 corresponds to the polynucleotide sequence of VEGFX10 illustrated in Figure 3.

Sequence ID No 14 corresponds to the polynucleotide sequence of VEGFX11 illustrated in Figure 4.

Sequence ID No 15 corresponds to the polynucleotide sequence of VEGFX12 illustrated in Figure 4.

Sequence ID No 16 corresponds to the polynucleotide sequence of VEGFX13 illustrated in Figure 4.

Sequence ID No 17 corresponds to the polynucleotide sequence of VEGFX14 illustrated in Figure 4.

Sequence ID No 18 corresponds to the polynucleotide sequence 5'-1 in Figure 8.

Sequence ID No 19 corresponds to the polynucleotide sequence 5'-2 in Figure 8.

Sequence ID No 20 corresponds to the polynucleotide sequence of VEGFX6 illustrated in Figure 13.

Sequence ID No 21 corresponds to the polynucleotide sequence of VEGFX7 illustrated in Figure 13.

Sequence ID No 22 corresponds to the polynucleotide sequence of VEGFX8 illustrated in Figure 13.

Sequence ID No 23 corresponds to the polynucleotide sequence of VEGFX9 illustrated in Figure 13.

Sequence ID No 24 corresponds to the polynucleotide sequence of VEGBAC1 illustrated in Figure 13.

Sequence ID No 25 corresponds to the polynucleotide sequence of VEGBAC2 illustrated in Figure 13.

Sequence ID No 26 corresponds to a polypeptide having the amino acid sequence from amino acid position 40 to 150 of the sequence of Figure 10.

Sequence ID No 27 corresponds to a polypeptide having the amino acid sequence illustrated in Figure 26.

Sequence ID No 28 corresponds to the sequence from

		position 5 to 508 of the nucleotide sequence illustrated in Figure 26.
5	Sequence ID No 29	corresponds to the nucleotide sequence from position 5 to 508 of the nucleotide sequence illustrated in Figure 26.
10	Sequence ID No 30	corresponds to the sequence from position 214 to 345 of the nucleotide sequence illustrated in Figure 10.

CLAIMS

1. A nucleic acid molecule encoding a VEGF-X protein or a functional equivalent, derivative or bioprecursor thereof, said protein comprising any of the sequences from position 23 to 345 of the amino acid sequence illustrated in Figure 10, or the complete sequence as illustrated in Figure 10.

2. A nucleic acid molecule according to claim 1 wherein said nucleic acid is a DNA molecule.

3. A nucleic acid molecule according to claim 1 or 2 wherein said nucleic acid is a cDNA molecule.

4. A nucleic acid molecule according to claim 3 comprising the nucleotide sequence from position 257 to 1291 of the nucleotide sequence illustrated in Figure 9, or sequences that hybridise thereto under high stringency conditions or the complement thereto.

5. An antisense molecule capable of hybridising to a molecule according to any of claims 1 to 4 under high stringency conditions.

6. A nucleic acid molecule according to any of claims 1 to 4 which is of mammalian origin.

7. A nucleic acid molecule according to claim 6 which is of human origin.

8. An isolated VEGF-X protein, or a functional equivalent, derivative or bioprecursor thereof, having an amino acid sequence from position 23 to 345 of the amino acid sequence illustrated in Figure 10 or the complete amino acid sequence of Figure 10.

9. A VEGF-X protein, or a functional equivalent, derivative or bioprecusor thereof, encoded by a nucleic acid molecule as defined in any of claims 1 to 4.

10. A protein according to claim 9, which comprises the amino acid sequence illustrated in Figure 10.

11. An expression vector comprising a nucleic acid molecule according to any of claims 1 to 4.

12. An expression vector according to claim 11 further comprising a nucleotide sequence encoding a reporter molecule.

13. An expression vector comprising an antisense molecule according to claim 5.

14. A nucleic acid molecule according to any of claims 1 to 4 or an antisense molecule according to claim 5 for use as a medicament.

15. A host cell transformed or transfected with an expression vector according to claim 11 or 12.

16. A host cell transformed or transfected with an expression vector according to claim 13.

17. A transgenic cell, tissue or organism comprising a transgene capable of expressing a VEGF-X protein according to claim 8 or 9.

18. A transgenic cell, tissue or organism according to claim 17, wherein said transgene is included in an expression vector.

19. A VEGF-X protein or a functional equivalent,

derivative or bioprecursor thereof, expressed by a cell according to claim 15.

20. A VEGF-X protein, or a functional equivalent, derivative or bioprecursor thereof, expressed by a transgenic cell, tissue or organism according to claim 17.

21. A process for producing a VEGF-X protein according to any of claims 8 to 10, said process comprising transforming a host cell or organism with an expression vector according to claim 11, and recovering the expressed protein from said host cell or organism.

22. An antibody capable of binding to a protein according to any of claims 8 to 10, or an epitope thereof.

23. An antibody according to claim 22 for use as a medicament.

24. A pharmaceutical composition comprising an antibody according to claim 22 together with a pharmaceutically acceptable carrier diluent or excipient thereof.

25. A method of identifying VEGF-X protein in a sample which method comprises contacting said sample with an antibody according to claim 22 and monitoring for binding of any protein to said antibody.

26. A kit for identifying the presence of VEGF-X protein in a sample which comprises an antibody according to claim 22 and means for contacting said antibody with said sample.

27. A method of identifying compounds which modulate angiogenesis which method comprises providing a host cell or organism according to claim 15 or a transgenic cell, tissue or organism according to claim 17, contacting a test compound with said cell, tissue or organism and monitoring for an effect of said compound on said VEGF compared to a host cell or organism according to claim 15 or a transgenic cell tissue or organism according to claim 17 which has not been contacted with said compound.

28. A compound identifiable according to the method of claim 27.

29. A compound according to claim 28 for use as a medicament.

30. A nucleic acid sequence comprising the nucleotide sequences illustrated in any of Figures 3, 5, 8 or 13.

31. A method for producing a polypeptide, said method comprising the steps of:

a) culturing the host cell of claim 15 under conditions suitable for expression of the polypeptide; and
b) recovering the polypeptide from the host cell culture.

32. A method of inhibiting angiogenic activity and inappropriate vascularisation including formation and proliferation of new blood vessels, growth and development of tissues, tissue regeneration and organ and tissue repair in a subject said method comprising administering to said subject an amount of an antisense molecule according to claim 5 in sufficient

concentration to reduce or prevent said angiogenic activity.

33. A method of inhibiting angiogenic activity or inappropriate vascularisation including any of formation and proliferation of new blood vessels, growth and development of tissues, tissue regeneration and organ and tissue repair in a subject said method comprising administering to said subject an amount of an antibody according to claim 22 in sufficient concentration to reduce or prevent said angiogenic activity or inappropriate vascularisation.

34. A method of inhibiting angiogenic activity or inappropriate vascularisation including any of formation and proliferation of new blood vessels, growth and development of tissues, tissue regeneration and organ and tissue repair in a subject, said method comprising implanting in said subject cells that express an antibody according to claim 22.

35. A method of treating or preventing any of cancer, rheumatoid arthritis, psoriasis and diabetic retinopathy, said method comprising administering to said subject an amount of an antisense molecule according to claim 5 in sufficient concentration to treat or prevent said disorders.

36. A method of treating or preventing any of cancer, rheumatoid arthritis, psoriasis and diabetic retinopathy, said method comprising administering to said subject an amount of an antibody according to claim 22 in sufficient concentration to reduce or prevent said disorders.

37. A method of promoting angiogenic activity or

vascularisation to promote wound healing, skin graft growth, tissue repair, proliferation of new blood vessels, tissue regeneration and organ repair which method comprises applying or delivering to a site of interest a therapeutically effective amount of any of a group selected from a protein according to claim 8 and a nucleic acid molecule encoding a VEGF-X protein or a functional equivalent, derivative or bioprecursor thereof comprising an amino acid sequence illustrated in Figure 10, an expression vector comprising said nucleic acid molecule and a pharmaceutical composition comprising any of said nucleic acid molecule and said protein.

38. A method of treating wounds selected from the group consisting of dermal ulcers, pressure sores, venous sores, diabetic ulcers and burns by applying to said wound a therapeutically effective amount of any of a VEGF-X protein according to claim 8, a pharmaceutical composition comprising said protein and a pharmaceutically acceptable carrier, diluent or excipient therefor.

39. A nucleic acid molecule encoding a polypeptide having a CUB domain said polypeptide comprising the amino acid sequence from position 40 to 150 of the sequence of Figure 10.

40. A nucleic acid molecule encoding a polypeptide having a CUB domain, said polypeptide comprising the amino acid sequence of Figure 26.

41. A nucleic acid molecule according to claim 39 or 40, comprising the nucleotide sequence from position 5 to 508 of the sequence illustrated in Figure 26.

42. A nucleic acid molecule according to any of

claims 39 to 41 comprising the nucleotide sequence illustrated in Figure 26.

43. A nucleic acid molecule encoding a VEGF like domain comprising the sequence from position 214-345 of the sequence of Figure 10 or the sequence from position 15 to 461 illustrated in Figure 24.

44. An expression vector comprising a nucleic acid molecule according to any of claims 39 to 42.

45. An expression vector comprising a nucleic acid molecule according to claim 43.

46. A host cell transformed or transfected with an expression vector according to claim 44.

47. A host cell transformed or transfected with an expression vector according to claim 45.

48. A protein expressed by the cell according to claim 46.

49. A protein expressed by the cell according to claim 47.

50. A method of identifying compounds that inhibit or enhance angiogenic activity, said method comprising contacting a cell expressing a VEGF receptor and/or a neuropilin 1 or 2 type receptor with said compound in the presence of a VEGF-X protein according to claim 8 and monitoring for the effect of said compound or said cell when compared to a cell which has not been contacted with said compound.

51. A compound identifiable according to the method

of claim 50 as an inhibitor or enhancer of angiogenic activity.

52. A method of inhibiting angiogenic activity or inappropriate vascularisation, said method comprising contacting a cell expressing a VEGF receptor and a neuropilin type receptor with a protein selected from any of a protein according to any of claims 8 to 10 and a protein according to claim 48 or a protein according to claim 49.

53. Use of a nucleotide sequence illustrated in any of Figures 14 and 15 in identifying a VEGF-X protein according to claim 8.

54. A nucleic acid molecule encoding a polypeptide comprising a CUB domain having the sequence from position 40 to 150 of the sequence of Figure 10 or from position 5 to 508 of the sequence of Figure 26 and a sequence encoding a VEGF domain.

55. A nucleic acid molecule according to claim 54 wherein said sequence encoding said VEGF domain is selected from the sequences encoding any of VEGF A to D or isoforms or variants thereof.

56. A nucleic acid molecule encoding a polypeptide comprising the amino acid sequence from position 40 to 150 of the sequence illustrated in Figure 10 for use as a medicament.

57. Use of a nucleic acid molecule encoding a polypeptide having the amino acid sequence from position 40 to 150 of the sequence illustrated in Figure 10 in the manufacture of a medicament for treatment of disease conditions associated with inappropriate angiogenesis such as tumour or cancer

growth, retinopathy, osteoarthritis or psoriasis.

58. A polypeptide comprising the amino acid sequence from position 40 to 150 of the sequence illustrated in figure 10 for use as a medicament.

59. A polypeptide comprising the amino acid sequence from position 40 to 150 of the sequence illustrated in Figure 10 in the manufacture of a medicament for the treatment of disease conditions associated with inappropriate angiogenesis such as tumour growth, retinopathy, osteoarthritis or psoriasis.

60. Use of a CUB domain comprising the amino acid sequence from position 40 to 150 of the sequence of Figure 10, or the amino acid sequence of Figure 26, to identify compounds which inhibit angiogenic activity in a method according to claim 50.

61. A method of inhibiting angiogenic activity and inappropriate vascularisation including formation and proliferation of new blood vessels, growth and development of tissues, tissue regeneration and organ and tissue repair in a subject said method comprising administering to said subject an amount of a polypeptide having an amino acid sequence from position 40 to 150 of the sequence illustrated in Figure 10 or a nucleic acid molecule according to any of claims 39 to 42 in sufficient concentration to reduce or prevent said angiogenic activity.

62. A method of treating or preventing any of cancer, rheumatoid arthritis, psoriasis and diabetic retinopathy, said method comprising administering to said subject an amount of a polypeptide having an amino acid sequence from position 40 to 150 of the sequence illustrated in Figure 10 or a nucleic acid

molecule according to any of claims 39 to 42 in sufficient concentration to treat or prevent said disorders.

63. An antisense molecule capable of hybridising to a molecule according to any of claims 39 to 42 under high stringency conditions.

64. An antisense molecule capable of hybridising to a molecule according to claim 43 under high stringency conditions.

65. A transgenic cell, tissue or organism comprising a transgene capable of expressing a protein according to claim 48.

66. A transgenic cell, tissue or organism comprising a transgene capable of expressing a protein according to claim 49.

67. A transgenic, cell tissue or organism according to claim 65 or 66, wherein said transgene is included in an expression vector according to claim 41 or 42.

68. An antibody capable of binding to a protein according to claim 48 or an epitope thereof.

69. An antibody capable of binding to a protein according to claim 49 or an epitope thereof.

70. A pharmaceutical composition comprising an antibody according to claim 68 or 69 together with a pharmaceutically acceptable carrier diluent or excipient therefor.

71. A pharmaceutical composition comprising a

compound according to claim 48 together with a pharmaceutically acceptable carrier, diluent or excipient therefor.

72. A nucleic acid molecule encoding a variant of a VEGF-X protein having any of the sequences of nucleotides illustrated in Figure 12.

FIG. 1

```
   1  AAAATGTATG GATACAACTT ACGTTTGATG AAAGATTTGG GCTTGAAGAC CCAGAAGATG
      TTTTACATAC CTATGTTGAA TGCAAACTAC TTTCTAAACC CGAACTTCTG GGTCTTCTAC

  61  ACATATGCAA GTATGATTTT GTAGAAGTTG AGGAACCCAG TGATGGAACT ATATTAGGGC
      TGTATACGTT CATACTAAAA CATCTTCAAC TCCTTGGGTC ACTACCTTGA TATAATCCCG

 121  GCTGGTGTGG TTCTGGTACT GTACCAGGAA AACAGATTTC TAAAGGAAAT CAAATTAGGA
      CGACCACACC AAGACCATGA CATGGTCCTT TTGTCTAAAG ATTTCCTTTA GTTTAATCCT

  +1                  MetAsn IlePheLeu LeuAsnLeuLeu ThrGluGlu ValArgLeu
                      ]----------------------------------------------
 181  TAAGATTTGT ATCTGATGAA TATTTTCCTT CTGAACCTTC TAACAGAGGA GGTAAGATTA
      ATTCTAAACA TAGACTACTT ATAAAAGGAA GACTTGGAAG ATTGTCTCCT CCATTCTAAT

  +1 TyrSerCysThr ProArgAsn PheSerVal SerIleArgGlu GluLeuLys ArgThrAsp
     ---------------------------------------------------------------
 241  TACAGCTGCA CACCTCGTAA CTTCTCAGTG TCCATAAGGG AAGAACTAAA GAGAACCGAT
      ATGTCGACGT GTGGAGCATT GAAGAGTCAC AGGTATTCCC TTCTTGATTT CTCTTGGCTA

  +1 ThrIlePheTrp ProGlyCys LeuLeuVal LysArgCysGly GlyAsnCys AlaCysCys
     ---------------------------------------------------------------
 301  ACCATTTTCT GGCCAGGTTG TCTCCTGGTT AAACGCTGTG GTGGGAACTG TGCCTGTTGT
      TGGTAAAAGA CCGGTCCAAC AGAGGACCAA TTTGCGACAC CACCCTTGAC ACGGACAACA

  +1 LeuHisAsnCys AsnGluCys GlnCysVal ProSerLysVal ThrLysLys TyrHisGlu
     ---------------------------------------------------------------
 361  CTCCACAATT GCAATGAATG TCAATGTGTC CCAAGCAAAG TTACTAAAAA ATACCACGAG
      GAGGTGTTAA CGTTACTTAC AGTTACACAG GGTTCGTTTC AATGATTTTT TATGGTGCTC

  +1 ValLeuGlnLeu ArgProLys ThrGlyVal ArgGlyLeuHis LysSerLeu ThrAspVal
     ---------------------------------------------------------------
 421  GTCCTTCAGT TGAGACCAAA GACCGGTGTC AGGGGATTGC ACAAATCACT CACCGACGTG
      CAGGAAGTCA ACTCTGGTTT CTGGCCACAG TCCCCTAACG TGTTTAGTGA GTGGCTGCAC

  +1 AlaLeuGluHis HisGluGlu CysAspCys ValCysArgGly SerThrGly Gly
     ------------------------------------------------------->
 481  GCCCTGGAGC ACCATGAGGA GTGTGACTGT GTGTGCAGAG GGAGCACAGG AGGATAGCCG
      CGGGACCTCG TGGTACTCCT CACACTGACA CACACGTCTC CCTCGTGTCC TCCTATCGGC

 541  CATCACCACC AGCAGCTCTT GCCCAGAGCT GTGCAGTGCA GTGGCTGATT CTATTAGAGA
      GTAGTGGTGG TCGTCGAGAA CGGGTCTCGA CACGTCACGT CACCGACTAA GATAATCTCT

 601  ACGTATGCGT TATCTCCATC CTTAATCTCA GTTGTTTGCT TCAAGGACCT TTCATCTTCA
      TGCATACGCA ATAGAGGTAG GAATTAGAGT CAACAAACGA AGTTCCTGGA AAGTAGAAGT

 661  GGATTTACAG TGCATTCTGA AAGAGGAGAC ATCAAACAGA ATTAGGAGTT GTGCAACAGC
      CCTAAATGTC ACGTAAGACT TTCTCCTCTG TAGTTTGTCT TAATCCTCAA CACGTTGTCG

 721  TCTTTTGAGA GGAGGCCTAA AGGACAGGAG AAAAGGTCTT CAATCGTGGA AAGAAAATTA
      AGAAAACTCT CCTCCGGATT TCCTGTCCTC TTTTCCAGAA GTTAGCACCT TTCTTTTAAT

 781  AATGTTGTAT TAAATAGATC ACCAGCTAGT TTCAGAGTTA CCATGTACGT ATTCCACTAG
      TTACAACATA ATTTATCTAG TGGTCGATCA AAGTCTCAAT GGTACATGCA TAAGGTGATC
```

FIG. 1 (CONTINUED).

```
841  CTGGGTTCTG TATTTCAGTT CTTTCGATAC GGCTTAGGGT AATGTCAGTA CAGGAAAAAA
     GACCCAAGAC ATAAAGTCAA GAAAGCTATG CCGAATCCCA TTACAGTCAT GTCCTTTTTT

901  ACTGTGCAAG TGAGCACCTG ATTCCGTTGC CTTGCTTAAC TCTAAAGCTC CATGTCCTGG
     TGACACGTTC ACTCGTGGAC TAAGGCAACG GAACGAATTG AGATTTCGAG GTACAGGACC

961  GCCTAAAATC GTATAAAATC TGGATTTTTT TTTTTTTTTT TGCTCATATT CACATATGTA
     CGGATTTTAG CATATTTTAG ACCTAAAAAA AAAAAAAAAA ACGAGTATAA GTGTATACAT

1021 AACCAGAACA TTCTATGTAC TACAAACCTG GTTTTTAAAA AGGAACTATG TTGCTATGAA
     TTGGTCTTGT AAGATACATG ATGTTTGGAC CAAAAATTTT TCCTTGATAC AACGATACTT

1081 TTAAACTTGT GTCGTGCTGA TAGGACAGAC TGGATTTTTC ATATTTCTTA TTAAAATTTC
     AATTTGAACA CAGCACGACT ATCCTGTCTG ACCTAAAAAG TATAAAGAAT AATTTTAAAG

1141 TGCCATTTAG AAGAAGAGAA CTACATTCAT GGTTTGGAAG AGATAAACCT GAAAAGAAGA
     ACGGTAAATC TTCTTCTCTT GATGTAAGTA CCAAACCTTC TCTATTTGGA CTTTTCTTCT

1201 GTGGCCTTAT CTTCACTTTA TCGATAAGTC AGTTTATTTG TTTCATTGTG TACATTTTTA
     CACCGGAATA GAAGTGAAAT AGCTATTCAG TCAAATAAAC AAAGTAACAC ATGTAAAAAT

1261 TATTCTCCTT TTGACATTAT AACTGTTGGC TTTTCTAATC TTGTTAAATA TATCTATTTT
     ATAAGAGGAA AACTGTAATA TTGACAACCG AAAAGATTAG AACAATTTAT ATAGATAAAA

1321 TACCAAAGGT ATTTAATATT CTTTTTTATG ACAACTTAGA TCAACTATTT TTAGCTTGGT
     ATGGTTTCCA TAAATTATAA GAAAAAATAC TGTTGAATCT AGTTGATAAA AATCGAACCA

1381 AAATTTTTCT AAACACAATT GTTATAGCCA GAGGAACAAA GATGATATAA AATATTGTTG
     TTTAAAAAGA TTTGTGTTAA CAATATCGGT CTCCTTGTTT CTACTATATT TTATAACAAC

1441 CTCTGACAAA AATACATGTA TTTCATTCTC GTATGGTGCT AGAGTTAGAT TAATCTGCAT
     GAGACTGTTT TTATGTACAT AAAGTAAGAG CATACCACGA TCTCAATCTA ATTAGACGTA

1501 TTTAAAAAAC TGAATTGGAA TAGAATTGGT AAGTTGCAAA GACTTTTTGA AAATAATTAA
     AAATTTTTTG ACTTAACCTT ATCTTAACCA TTCAACGTTT CTGAAAAACT TTTATTAATT

1561 ATTATCATAT CTTCCATTCC TGTTATTGGA GATGAAAATA AAAAGCAACT TATGAAAGTA
     TAATAGTATA GAAGGTAAGG ACAATAACCT CTACTTTTAT TTTTCGTTGA ATACTTTCAT

1621 GACATTCAGA TCCAGCCATT ACTAACCTAT TCCTTTTTTG GGGAAATCTG AGCCTAGCTC
     CTGTAAGTCT AGGTCGGTAA TGATTGGATA AGGAAAAAAC CCCTTTAGAC TCGGATCGAG

1681 AGAAAAACAT AAAGCACCTT GAAAAAGACT TGGCAGCTTC CTGATAAAGC GTGCTGTGCT
     TCTTTTTGTA TTTCGTGGAA CTTTTTCTGA ACCGTCGAAG GACTATTTCG CACGACACGA

1741 GTGCAGTAGG AACACATCCT ATTTATTGTG ATGTTGTGGT TTTATTATCT TAAACTCTGT
     CACGTCATCC TTGTGTAGGA TAAATAACAC TACAACACCA AAATAATAGA ATTTGAGACA

1801 TCCATACACT TGTATAAATA CATGGATATT TTTATGTACA GAAGTATGTC TCTTAACCAG
     AGGTATGTGA ACATATTTAT GTACCTATAA AAATACATGT CTTCATACAG AGAATTGGTC

1861 TTCACTTATT GTACCTGG
     AAGTGAATAA CATGGACC
```

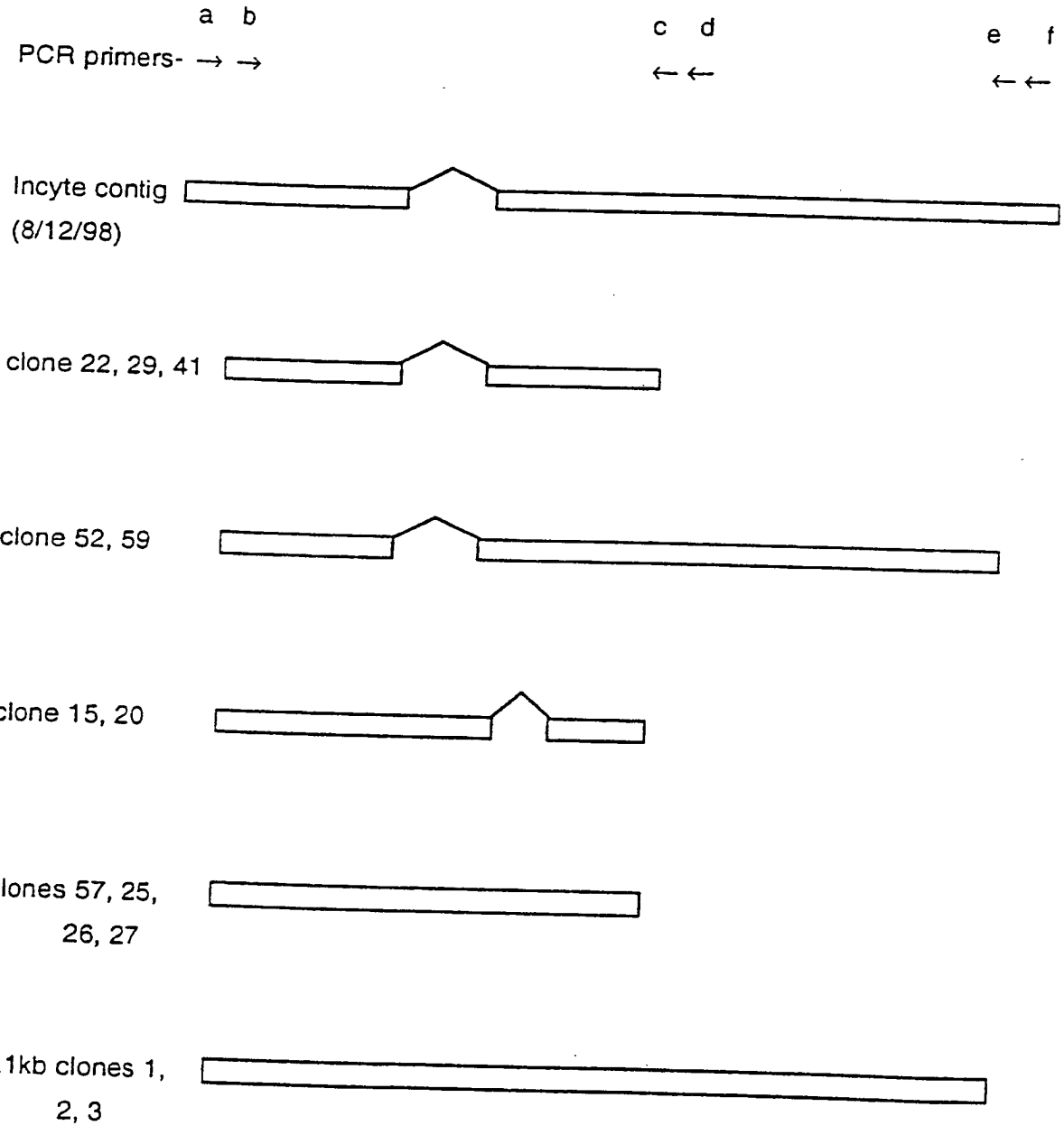
FIG. 2. **Predicted VEGF-like protein encoded by Incyte contig of 8/12/98**

```
  1  MNIFLLNLLT EEVRLYSCTP RNFSVSIREE LKRTDTIFWP GCLLVKRCGG
 51  NCACCLHNCN ECQCVPSKVT KKYHEVLQLR PKTGVRGLHK SLTDVALEHH
101  EECDCVCRGS TGG
```

FIG. 3. **PCR primers for cloning VEGF-X**

vegfX1	AAAATGTATGGATACAACTTAC
vegfX2	GTTTGATGAAAGATTTGGGCTTG
vegfX3	TTTCTAAAGGAAATCAAATTAG
vegfX4	GATAAGATTTGTATCTGATG
vegfX5	GATGTCTCCTCTTTCAG
vegfX6	GCACAACTCCTAATTCTG
vegfX7	AGCACCTGATTCCGTTGC
vegfX8	TAGTACATAGAATGTTCTGG
vegfX9	AAGAGACATACTTCTGTAC
vegfX10	CCAGGTACAATAAGTGAACTG

FIG 4. **Variants isolated by PCR (at 8/2/99, all cloned and sequenced at JRF)**



primers- (see fig 3) a- vegfX1 b- vegfX2 c- vegfX5 d- vegfX6 e- vegfX9 f- vegfX10

FIG. 5. **VEGF-X 5' RACE primers**

vegfX11	CCTTTAGAAATCTGTTTTCCTGGTACAG
vegfX12	GGAAAATATTCATCAGATACAAATCTTATCC
vegfX13	GGTCCAGTGGCAAAGCTGAAGG
vegfX14	CTGGTTCAAGATATCGAATAAGGTCTTCC

FIG. 6. DNA sequence assembled from in-house clones and 5'RACE

```
   1  TGCCAGAGCA GGTGGGCGCT TCCACCCCAG TGCAGCCTTC CCCTGGCGGT GGTGAAAGAG
      ACGGTCTCGT CCACCCGCGA AGGTGGGGTC ACGTCGGAAG GGGACCGCCA CCACTTTCTC

  61  ACTCGGGAGT CGCTGCTTCC AAAGTGCCCG CCGTGAGTGA GCTCTCACCC CAGTCAGCCA
      TGAGCCCTCA GCGACGAAGG TTTCACGGGC GGCACTCACT CGAGAGTGGG GTCAGTCGGT

  +2   MetSerLeu PheGlyLeuLeu LeuLeuThr SerAlaLeu AlaGlyGlnArg GlnGlyTh
       ]----------------------------------------------------------------
 121  AATGAGCCTC TTCGGGCTTC TCCTGCTGAC ATCTGCCCTG GCCGGCCAGA GACAGGGGAC
      TTACTCGGAG AAGCCCGAAG AGGACGACTG TAGACGGGAC CGGCCGGTCT CTGTCCCCTG

  +2  rGlnAlaGlu SerAsnLeuSer SerLysPhe GlnPheSer SerAsnLysGlu GlnAsnGl
      ----------------------------------------------------------------
 181  TCAGGCGGAA TCCAACCTGA GTAGTAAATT CCAGTTTTCC AGCAACAAGG AACAGAACGG
      AGTCCGCCTT AGGTTGGACT CATCATTTAA GGTCAAAAGG TCGTTGTTCC TTGTCTTGCC

  +2  yValGlnAsp ProGlnHisGlu ArgIleIle ThrValSer ThrAsnGlySer IleHisSe
      ----------------------------------------------------------------
 241  AGTACAAGAT CCTCAGCATG AGAGAATTAT TACTGTGTCT ACTAATGGAA GTATTCACAG
      TCATGTTCTA GGAGTCGTAC TCTCTTAATA ATGACACAGA TGATTACCTT CATAAGTGTC

  +2  rProArgPhe ProHisThrTyr ProArgAsn ThrValLeu ValTrpArgLeu ValAlaVa
      ----------------------------------------------------------------
 301  CCCAAGGTTT CCTCATACTT ATCCAAGAAA TACGGTCTTG GTATGGAGAT TAGTAGCAGT
      GGGTTCCAAA GGAGTATGAA TAGGTTCTTT ATGCCAGAAC CATACCTCTA ATCATCGTCA

  +2  lGluGluAsn ValTrpIleGln LeuThrPhe AspGluArg PheGlyLeuGlu AspProGl
      ----------------------------------------------------------------
 361  AGAGGAAAAT GTATGGATAC AACTTACGTT TGATGAAAGA TTTGGGCTTG AAGACCCAGA
      TCTCCTTTTA CATACCTATG TTGAATGCAA ACTACTTTCT AAACCCGAAC TTCTGGGTCT

  +2  uAspAspIle CysLysTyrAsp PheValGlu ValGluGlu ProSerAspGly ThrIleLe
      ----------------------------------------------------------------
 421  AGATGACATA TGCAAGTATG ATTTTGTAGA AGTTGAGGAA CCCAGTGATG GAACTATATT
      TCTACTGTAT ACGTTCATAC TAAAACATCT TCAACTCCTT GGGTCACTAC CTTGATATAA

  +2  uGlyArgTrp CysGlySerGly ThrValPro GlyLysGln IleSerLysGly AsnGlnIl
      ----------------------------------------------------------------
 481  AGGGCGCTGG TGTGGTTCTG GTACTGTACC AGGAAAACAG ATTTCTAAAG GAAATCAAAT
      TCCCGCGACC ACACCAAGAC CATGACATGG TCCTTTTGTC TAAAGATTTC CTTTAGTTTA

  +2  eArgIleArg PheValSerAsp GluTyrPhe ProSerGlu ProGlyPheCys IleHisTy
      ----------------------------------------------------------------
 541  TAGGATAAGA TTTGTATCTG ATGAATATTT TCCTTCTGAA CCAGGGTTCT GCATCCACTA
      ATCCTATTCT AAACATAGAC TACTTATAAA AGGAAGACTT GGTCCCAAGA CGTAGGTGAT

  +2  rAsnIleVal MetProGlnPhe ThrGluAla ValSerPro SerValLeuPro ProSerAl
      ----------------------------------------------------------------
 601  CAACATTGTC ATGCCACAAT TCACAGAAGC TGTGAGTCCT TCAGTGCTAC CCCCTTCAGC
      GTTGTAACAG TACGGTGTTA AGTGTCTTCG ACACTCAGGA AGTCACGATG GGGGAAGTCG

  +2  aLeuProLeu AspLeuLeuAsn AsnAlaIle ThrAlaPhe SerThrLeuGlu AspLeuIl
      ----------------------------------------------------------------
 661  TTTGCCACTG GACCTGCTTA ATAATGCTAT AACTGCCTTT AGTACCTTGG AAGACCTTAT
```

FIG. 6 (CONTINUED 1).

```
   +2 eArgTyrLeu GluProGluArg TrpGlnLeu AspLeuGlu AspLeuTyrArg ProThrTr
      ------------------------------------------------------------------
  721 TCGATATCTT GAACCAGAGA GATGGCAGTT GGACTTAGAA GATCTATATA GGCCAACTTG
      AGCTATAGAA CTTGGTCTCT CTACCGTCAA CCTGAATCTT CTAGATATAT CCGGTTGAAC

   +2 pGlnLeuLeu GlyLysAlaPhe ValPheGly ArgLysSer ArgValValAsp LeuAsnLe
      ------------------------------------------------------------------
  781 GCAACTTCTT GGCAAGGCTT TTGTTTTTGG AAGAAAATCC AGAGTGGTGG ATCTGAACCT
      CGTTGAAGAA CCGTTCCGAA AACAAAAACC TTCTTTTAGG TCTCACCACC TAGACTTGGA

   +2 uLeuThrGlu GluValArgLeu TyrSerCys ThrProArg AsnPheSerVal SerIleAr
      ------------------------------------------------------------------
  841 TCTAACAGAG GAGGTAAGAT TATACAGCTG CACACCTCGT AACTTCTCAG TGTCCATAAG
      AGATTGTCTC CTCCATTCTA ATATGTCGAC GTGTGGAGCA TTGAAGAGTC ACAGGTATTC

   +2 gGluGluLeu LysArgThrAsp ThrIlePhe TrpProGly CysLeuLeuVal LysArgCy
      ------------------------------------------------------------------
  901 GGAAGAACTA AAGAGAACCG ATACCATTTT CTGGCCAGGT TGTCTCCTGG TTAAACGCTG
      CCTTCTTGAT TTCTCTTGGC TATGGTAAAA GACCGGTCCA ACAGAGGACC AATTTGCGAC

   +2 sGlyGlyAsn CysAlaCysCys LeuHisAsn CysAsnGlu CysGlnCysVal ProSerLy
      ------------------------------------------------------------------
  961 TGGTGGGAAC TGTGCCTGTT GTCTCCACAA TTGCAATGAA TGTCAATGTG TCCCAAGCAA
      ACCACCCTTG ACACGGACAA CAGAGGTGTT AACGTTACTT ACAGTTACAC AGGGTTCGTT

   +2 sValThrLys LysTyrHisGlu ValLeuGln LeuArgPro LysThrGlyVal ArgGlyLe
      ------------------------------------------------------------------
 1021 AGTTACTAAA AAATACCACG AGGTCCTTCA GTTGAGACCA AAGACCGGTG TCAGGGGATT
      TCAATGATTT TTTATGGTGC TCCAGGAAGT CAACTCTGGT TTCTGGCCAC AGTCCCCTAA

   +2 uHisLysSer LeuThrAspVal AlaLeuGlu HisHisGlu GluCysAspCys ValCysAr
      ------------------------------------------------------------------
 1081 GCACAAATCA CTCACCGACG TGGCCCTGGA GCACCATGAG GAGTGTGACT GTGTGTGCAG
      CGTGTTTAGT GAGTGGCTGC ACCGGGACCT CGTGGTACTC CTCACACTGA CACACACGTC

   +2 gGlySerThr GlyGly
      ---------------->
 1141 AGGGAGCACA GGAGGATAGC CGCATCACCA CCAGCAGCTC TTGCCCAGAG CTGTGCAGTG
      TCCCTCGTGT CCTCCTATCG GCGTAGTGGT GGTCGTCGAG AACGGGTCTC GACACGTCAC

 1201 CAGTGGCTGA TTCTATTAGA GAACGTATGC GTTATCTCCA TCCTTAATCT CAGTTGTTTG
      GTCACCGACT AAGATAATCT CTTGCATACG CAATAGAGGT AGGAATTAGA GTCAACAAAC

 1261 CTTCAAGGAC CTTTCATCTT CAGGATTTAC AGTGCATTCT GAAAGAGGAG ACATCAAACA
      GAAGTTCCTG GAAAGTAGAA GTCCTAAATG TCACGTAAGA CTTTCTCCTC TGTAGTTTGT

 1321 GAATTAGGAG TTGTGCAACA GCTCTTTTGA GAGGAGGCCT AAAGGACAGG AGAAAAGGTC
      CTTAATCCTC AACACGTTGT CGAGAAAACT CTCCTCCGGA TTTCCTGTCC TCTTTTCCAG

 1381 TTCAATCGTG GAAAGAAAAT TAAATGTTGT ATTAAATAGA TCACCAGCTA GTTTCAGAGT
      AAGTTAGCAC CTTTCTTTTA ATTTACAACA TAATTTATCT AGTGGTCGAT CAAAGTCTCA

 1441 TACCATGTAC GTATTCCACT AGCTGGGTTC TGTATTTCAG TTCTTTCGAT ACGGCTTAGG
      ATGGTACATG CATAAGGTGA TCGACCCAAG ACATAAAGTC AAGAAAGCTA TGCCGAATCC

 1501 GTAATGTCAG TACAGGAAAA AAACTGTGCA AGTGAGCACC TGATTCCGTT GCCTTGCTTA
```

FIG. 6 (CONTINUED 2).

```
1561  ACTCTAAAGC TCCATGTCCT GGGCCTAAAA TCGTATAAAA TCTGGATTTT TTTTTTTTTT
      TGAGATTTCG AGGTACAGGA CCCGGATTTT AGCATATTTT AGACCTAAAA AAAAAAAAAA

1621  TTTGCTCATA TTCACATATG TAAACCAGAA CATTCTATGT ACTACAAACC TGGTTTTTAA
      AAACGAGTAT AAGTGTATAC ATTTGGTCTT GTAAGATACA TGATGTTTGG ACCAAAAATT

1681  AAAGGAACTA TGTTGCTATG AATTAAACTT GTGTCGTGCT GATAGGACAG ACTGGATTTT
      TTTCCTTGAT ACAACGATAC TTAATTTGAA CACAGCACGA CTATCCTGTC TGACCTAAAA

1741  TCATATTTCT TATTAAAATT TCTGCCATTT AGAAGAAGAG AACTACATTC ATGGTTTGGA
      AGTATAAAGA ATAATTTTAA AGACGGTAAA TCTTCTTCTC TTGATGTAAG TACCAAACCT

1801  AGAGATAAAC CTGAAAAGAA GAGTGGCCTT ATCTTCACTT TATCGATAAG CCAGTTTATT
      TCTCTATTTG GACTTTTCTT CTCACCGGAA TAGAAGTGAA ATAGCTATTC GGTCAAATAA

1861  TGTTTCATTG TGTACATTTT TATATTCTCC TTTTGACATT ATAACTGTTG GCTTTTCTAA
      ACAAAGTAAC ACATGTAAAA ATATAAGAGG AAAACTGTAA TATTGACAAC CGAAAAGATT

1921  TCTTGTTAAA TATATCTATT TTTACCAAAG GTATTTAATA TTCTTTTTTA TGACAACTTA
      AGAACAATTT ATATAGATAA AAATGGTTTC CATAAATTAT AAGAAAAAAT ACTGTTGAAT

1981  GATCAACTAT TTTTAGCTTG GTAAATTTTT CTAAACACAA TTGTTATAGC CAGAGGAACA
      CTAGTTGATA AAAATCGAAC CATTTAAAAA GATTTGTGTT AACAATATCG GTCTCCTTGT

2041  AAGATGATAT AAAATATTGT TGCTCTGACA AAAATACATG TATTTCATTC TCGTATGGTG
      TTCTACTATA TTTTATAACA ACGAGACTGT TTTTATGTAC ATAAAGTAAG AGCATACCAC

2101  CTAGAGTTAG ATTAATCTGC ATTTTAAAAA ACTGAATTGG AATAGAATTG GTAAGTTGCA
      GATCTCAATC TAATTAGACG TAAAATTTTT TGACTTAACC TTATCTTAAC CATTCAACGT

2161  AAGACTTTTT GAAAATAATT AAATTATCAT ATCTTCCATT CCTGTTATTG GAGATGAAAA
      TTCTGAAAAA CTTTTATTAA TTTAATAGTA TAGAAGGTAA GGACAATAAC CTCTACTTTT

2221  TAAAAAGCAA CTTATGAAAG TAGACATTCA GATCCAGCCA TTACTAACCT ATTCCTTTTT
      ATTTTTCGTT GAATACTTTC ATCTGTAAGT CTAGGTCGGT AATGATTGGA TAAGGAAAAA

2281  TGGGGAAATC TGAGCCTAGC TCAGAAAAAC ATAAAGCACC TTGAAAAAGA CTTGGCAGCT
      ACCCCTTTAG ACTCGGATCG AGTCTTTTTG TATTTCGTGG AACTTTTTCT GAACCGTCGA

2341  TCCTGATAAA GCGTGCTGTG CTGTGCAGTA GGAACACATC CTATTTATTG TGATGTTGTG
      AGGACTATTT CGCACGACAC GACACGTCAT CCTTGTGTAG GATAAATAAC ACTACAACAC

2401  GTTTTATTAT CTTAAACTCT GTTCCATACA CTTGTATAAA TACATGGATA TTTTTATGTA
      CAAAATAATA GAATTTGAGA CAAGGTATGT GAACATATTT ATGTACCTAT AAAAATACAT

2461  CAGAAGTATG TCTCT
      GTCTTCATAC AGAGA
```

FIG. 7. New Sequence + Incyte ESTs

```
   1  ATTTGTTTAA ACCTTGGGAA ACTGGTTCAG GTCCAGGTTT TGCTTTGATC CTTTTCAAAA
      TAAACAAATT TGGAACCCTT TGACCAAGTC CAGGTCCAAA ACGAAACTAG GAAAAGTTTT

  61  ACTGGAGACA CAGAAGAGGG CTTCTAGGAA AAAGTTTTGG GATGGGATTA TGTGGAAACT
      TGACCTCTGT GTCTTCTCCC GAAGATCCTT TTTCAAAACC CTACCCTAAT ACACCTTTGA

 121  ACCCTGCGAT TCTCTGCTGC CAGAGCAGGC TCGGCGCTTC CACCCCAGTG CAGCCTTCCC
      TGGGACGCTA AGAGACGACG GTCTCGTCCG AGCCGCGAAG GTGGGGTCAC GTCGGAAGGG

 181  CTGGCGGTGG TGAAAGAGAC TCGGGAGTCG CTGCTTCCAA AGTGCCCGCC GTGAGTGAGC
      GACCGCCACC ACTTTCTCTG AGCCCTCAGC GACGAAGGTT TCACGGGCGG CACTCACTCG

  +2                      Met SerLeuPhe GlyLeuLeu LeuLeuThrSer AlaLeuAl
                          ]----------------------------------------
 241  TCTCACCCCA GTCAGCCAAA TGAGCCTCTT CGGGCTTCTC CTGCTGACAT CTGCCCTGGC
      AGAGTGGGGT CAGTCGGTTT ACTCGGAGAA GCCCGAAGAG GACGACTGTA GACGGGACCG

  +2 aGlyGlnArg GlnGlyThrGln AlaGluSer AsnLeuSer SerLysPheGln PheSerSe
     ----------------------------------------------------------------
 301  CGGCCAGAGA CAGGGGACTC AGGCGGAATC CAACCTGAGT AGTAAATTCC AGTTTTCCAG
      GCCGGTCTCT GTCCCCTGAG TCCGCCTTAG GTTGGACTCA TCATTTAAGG TCAAAAGGTC

  +2 rAsnLysGlu GlnTyrGlyVal GlnAspPro GlnHisGlu ArgIleIleThr ValSerTh
     ----------------------------------------------------------------
 361  CAACAAGGAA CAGTACGGAG TACAAGATCC TCAGCATGAG AGAATTATTA CTGTGTCTAC
      GTTGTTCCTT GTCATGCCTC ATGTTCTAGG AGTCGTACTC TCTTAATAAT GACACAGATG

  +2 rAsnGlySer IleHisSerPro ArgPhePro HisThrTyr ProArgAsnThr ValLeuVa
     ----------------------------------------------------------------
 421  TAATGGAAGT ATTCACAGCC CAAGGTTTCC TCATACTTAT CCAAGAAATA CGGTCTTGGT
      ATTACCTTCA TAAGTGTCGG GTTCCAAAGG AGTATGAATA GGTTCTTTAT GCCAGAACCA

  +2 lTrpArgLeu ValAlaValGlu GluAsnVal TrpIleGln LeuThrPheAsp GluArgPh
     ----------------------------------------------------------------
 481  ATGGAGATTA GTAGCAGTAG AGGAAAATGT ATGGATACAA CTTACGTTTG ATGAAAGATT
      TACCTCTAAT CATCGTCATC TCCTTTTACA TACCTATGTT GAATGCAAAC TACTTTCTAA

  +2 eGlyLeuGlu AspProGluAsp AspIleCys LysTyrAsp PheValGluVal GluGluPr
     ----------------------------------------------------------------
 541  TGGGCTTGAA GACCCAGAAG ATGACATATG CAAGTATGAT TTTGTAGAAG TTGAGGAACC
      ACCCGAACTT CTGGGTCTTC TACTGTATAC GTTCATACTA AAACATCTTC AACTCCTTGG

  +2 oSerAspGly ThrIleLeuGly ArgTrpCys GlySerGly ThrValProGly LysGlnIl
     ----------------------------------------------------------------
 601  CAGTGATGGA ACTATATTAG GGCGCTGGTG TGGTTCTGGT ACTGTACCAG GAAAACAGAT
      GTCACTACCT TGATATAATC CCGCGACCAC ACCAAGACCA TGACATGGTC CTTTTGTCTA

  +2 eSerLysGly AsnGlnIleArg IleArgPhe ValSerAsp GluTyrPhePro SerGluPr
     ----------------------------------------------------------------
 661  TTCTAAAGGA AATCAAATTA GGATAAGATT TGTATCTGAT GAATATTTTC CTTCTGAACC
      AAGATTTCCT TTAGTTTAAT CCTATTCTAA ACATAGACTA CTTATAAAAG GAAGACTTGG
```

FIG. 7 (CONTINUED 1).

```
   +2 oGlyPheCys IleHisTyrAsn IleValMet ProGlnPhe ThrGluAlaVal SerProSe
      ------------------------------------------------------------------
  721 AGGGTTCTGC ATCCACTACA ACATTGTCAT GCCACAATTC ACAGAAGCTG TGAGTCCTTC
      TCCCAAGACG TAGGTGATGT TGTAACAGTA CGGTGTTAAG TGTCTTCGAC ACTCAGGAAG

   +2 rValLeuPro ProSerAlaLeu ProLeuAsp LeuLeuAsn AsnAlaIleThr AlaPheSe
      ------------------------------------------------------------------
  781 AGTGCTACCC CCTTCAGCTT TGCCACTGGA CCTGCTTAAT AATGCTATAA CTGCCTTTAG
      TCACGATGGG GGAAGTCGAA ACGGTGACCT GGACGAATTA TTACGATATT GACGGAAATC

   +2 rThrLeuGlu AspLeuIleArg TyrLeuGlu ProGluArg TrpGlnLeuAsp LeuGluAs
      ------------------------------------------------------------------
  841 TACCTTGGAA GACCTTATTC GATATCTTGA ACCAGAGAGA TGGCAGTTGG ACTTAGAAGA
      ATGGAACCTT CTGGAATAAG CTATAGAACT TGGTCTCTCT ACCGTCAACC TGAATCTTCT

   +2 pLeuTyrArg ProThrTrpGln LeuLeuGly LysAlaPhe ValPheGlyArg LysSerAr
      ------------------------------------------------------------------
  901 TCTATATAGG CCAACTTGGC AACTTCTTGG CAAGGCTTTT GTTTTTGGAA GAAAATCCAG
      AGATATATCC GGTTGAACCG TTGAAGAACC GTTCCGAAAA CAAAAACCTT CTTTTAGGTC

   +2 gValValAsp LeuAsnLeuLeu ThrGluGlu ValArgLeu TyrSerCysThr ProArgAs
      ------------------------------------------------------------------
  961 AGTGGTGGAT CTGAACCTTC TAACAGAGGA GGTAAGATTA TACAGCTGCA CACCTCGTAA
      TCACCACCTA GACTTGGAAG ATTGTCTCCT CCATTCTAAT ATGTCGACGT GTGGAGCATT

   +2 nPheSerVal SerIleArgGlu GluLeuLys ArgThrAsp ThrIlePheTrp ProGlyCy
      ------------------------------------------------------------------
 1021 CTTCTCAGTG TCCATAAGGG AAGAACTAAA GAGAACCGAT ACCATTTTCT GGCCAGGTTG
      GAAGAGTCAC AGGTATTCCC TTCTTGATTT CTCTTGGCTA TGGTAAAAGA CCGGTCCAAC

   +2 sLeuLeuVal LysArgCysGly GlyAsnCys AlaCysCys LeuHisAsnCys AsnGluCy
      ------------------------------------------------------------------
 1081 TCTCCTGGTT AAACGCTGTG GTGGGAACTG TGCCTGTTGT CTCCACAATT GCAATGAATG
      AGAGGACCAA TTTGCGACAC CACCCTTGAC ACGGACAACA GAGGTGTTAA CGTTACTTAC

   +2 sGlnCysVal ProSerLysVal ThrLysLys TyrHisGlu ValLeuGlnLeu ArgProLy
      ------------------------------------------------------------------
 1141 TCAATGTGTC CCAAGCAAAG TTACTAAAAA ATACCACGAG GTCCTTCAGT TGAGACCAAA
      AGTTACACAG GGTTCGTTTC AATGATTTTT TATGGTGCTC CAGGAAGTCA ACTCTGGTTT

   +2 sThrGlyVal ArgGlyLeuHis LysSerLeu ThrAspVal AlaLeuGluHis HisGluGl
      ------------------------------------------------------------------
 1201 GACCGGTGTC AGGGGATTGC ACAAATCACT CACCGACGTG GCCCTGGAGC ACCATGAGGA
      CTGGCCACAG TCCCCTAACG TGTTTAGTGA GTGGCTGCAC CGGGACCTCG TGGTACTCCT

   +2 uCysAspCys ValCysArgGly SerThrGly Gly
      ------------------------------------>
 1261 GTGTGACTGT GTGTGCAGAG GGAGCACAGG AGGATAGCCG CATCACCACC AGCAGCTCTT
      CACACTGACA CACACGTCTC CCTCGTGTCC TCCTATCGGC GTAGTGGTGG TCGTCGAGAA

 1321 GCCCAGAGCT GTGCAGTGCA GTGGCTGATT CTATTAGAGA ACGTATGCGT TATCTCCATC
      CGGGTCTCGA CACGTCACGT CACCGACTAA GATAATCTCT TGCATACGCA ATAGAGGTAG

 1381 CTTAATCTCA GTTGTTTGCT TCAAGGACCT TTCATCTTCA GGATTTACAG TGCATTCTGA
      GAATTAGAGT CAACAAACGA AGTTCCTGGA AAGTAGAAGT CCTAAATGTC ACGTAAGACT
```

FIG. 7 (CONTINUED 2).

```
1441  AAGAGGAGAC ATCAAACAGA ATTAGGAGTT GTGCAACAGC TCTTTTGAGA GGAGGCCTAA
      TTCTCCTCTG TAGTTTGTCT TAATCCTCAA CACGTTGTCG AGAAAACTCT CCTCCGGATT

1501  AGGACAGGAG AAAAGGTCTT CAATCGTGGA AAGAAAATTA AATGTTGTAT TAAATAGATC
      TCCTGTCCTC TTTTCCAGAA GTTAGCACCT TTCTTTTAAT TTACAACATA ATTTATCTAG

1561  ACCAGCTAGT TTCAGAGTTA CCATGTACGT ATTCCACTAG CTGGGTTCTG TATTTCAGTT
      TGGTCGATCA AAGTCTCAAT GGTACATGCA TAAGGTGATC GACCCAAGAC ATAAAGTCAA

1621  CTTTCGATAC GGCTTAGGGT AATGTCAGTA CAGGAAAAAA ACTGTGCAAG TGAGCACCTG
      GAAAGCTATG CCGAATCCCA TTACAGTCAT GTCCTTTTTT TGACACGTTC ACTCGTGGAC

1681  ATTCCGTTGC CTTGGCTTAA CTCTAAAGCT CCATGTCCTG GGCCTAAAAT CGTATAAAAT
      TAAGGCAACG GAACCGAATT GAGATTTCGA GGTACAGGAC CCGGATTTTA GCATATTTTA

1741  CTGGATTTTT TTTTTTTTTT TTGCGCATAT TCACATATGT AAACCAGAAC ATTCTATGTA
      GACCTAAAAA AAAAAAAAAA AACGCGTATA AGTGTATACA TTTGGTCTTG TAAGATACAT

1801  CTACAAACCT GGTTTTTAAA AAGGAACTAT GTTGCTATGA ATTAAACTTG TGTCATGCTG
      GATGTTTGGA CCAAAAATTT TTCCTTGATA CAACGATACT TAATTTGAAC ACAGTACGAC

1861  ATAGGACAGA CTGGATTTTT CATATTTCTT ATTAAAATTT CTGCCATTTA GAAGAAGAGA
      TATCCTGTCT GACCTAAAAA GTATAAAGAA TAATTTTAAA GACGGTAAAT CTTCTTCTCT

1921  ACTACATTCA TGGTTTGGAA GAGATAAACC TGAAAAGAAG AGTGGCCTTA TCTTCACTTT
      TGATGTAAGT ACCAAACCTT CTCTATTTGG ACTTTTCTTC TCACCGGAAT AGAAGTGAAA

1981  ATCGATAAGT CAGTTTATTT GTTTCATTGT GTACATTTTT ATATTCTCCT TTTGACATTA
      TAGCTATTCA GTCAAATAAA CAAAGTAACA CATGTAAAAA TATAAGAGGA AAACTGTAAT

2041  TAACTGTTGG CTTTTCTAAT CTTGTTAAAT ATATCTATTT TTACCAAAGG TATTTAATAT
      ATTGACAACC GAAAAGATTA GAACAATTTA TATAGATAAA AATGGTTTCC ATAAATTATA

2101  TCTTTTTTAT GACAACTTAG ATCAACTATT TTTAGCTTGG TAAATTTTTC TAAACACAAT
      AGAAAAAATA CTGTTGAATC TAGTTGATAA AAATCGAACC ATTTAAAAAG ATTTGTGTTA

2161  TGTTATAGCC AGAGGAACAA AGATGATATA AAATATTGTT GCTCTGACAA AAATACATGT
      ACAATATCGG TCTCCTTGTT TCTACTATAT TTTATAACAA CGAGACTGTT TTTATGTACA

2221  ATTTCATTCT CGTATGGTGC TAGAGTTAGA TTAATCTGCA TTTTAAAAAA CTGAATTGGA
      TAAAGTAAGA GCATACCACG ATCTCAATCT AATTAGACGT AAAATTTTTT GACTTAACCT

2281  ATAGAATTGG TAAGTTGCAA AGACTTTTTG AAAATAATTA AATTATCATA TCTTCCATTC
      TATCTTAACC ATTCAACGTT TCTGAAAAAC TTTTATTAAT TTAATAGTAT AGAAGGTAAG

2341  CTGTTATTGG AGATGAAAAT AAAAAGCAAC TTATGAAAGT AGACATTCAG ATCCAGCCAT
      GACAATAACC TCTACTTTTA TTTTTCGTTG AATACTTTCA TCTGTAAGTC TAGGTCGGTA

2401  TACTAACCTA TTCCTTTTTT GGGGAAATCT GAGCCTAGCT CAGAAAAACA TAAAGCACCT
      ATGATTGGAT AAGGAAAAAA CCCCTTTAGA CTCGGATCGA GTCTTTTTGT ATTTCGTGGA

2461  TGAAAAAGAC TTGGCAGCTT CCTGATAAAG CGTGCTGTGC TGTGCAGTAG GAACACATCC
      ACTTTTTCTG AACCGTCGAA GGACTATTTC GCACGACACG ACACGTCATC CTTGTGTAGG

2521  TATTTATTGT GATGTTGTGG TTTTATTATC TTAAACTCTG TTCCATACAC TTGTATAAAT
      ATAAATAACA CTACAACACC AAAATAATAG AATTTGAGAC AAGGTATGTG AACATATTTA
```

FIG. 7 (CONTINUED 3).

```
2581  ACATGGATAT TTTTATGTAC AGAAGTATGT CTCTTAACCA GTTCACTTAT TGTACTCTGG
      TGTACCTATA AAAATACATG TCTTCATACA GAGAATTGGT CAAGTGAATA ACATGAGACC

2641  CAATTTAAAA GAAAATCAGT AAAATATTTT GCTTGTAAAA TGCTTAATAT CGTGCCTAGG
      GTTAAATTTT CTTTTAGTCA TTTTATAAAA CGAACATTTT ACGAATTATA GCACGGATCC

2701  TTATGTGGTG ACTATTTGAA TCAAAAATGT ATTGAATCAT CAAATAAAAG AATGTGGCTA
      AATACACCAC TGATAAACTT AGTTTTTACA TAACTTAGTA GTTTATTTTC TTACACCGAT

2761  TTTTGGGGAG AAAATT
      AAAACCCCTC TTTTAA
```

FIG. 8. **Additional oligonucleotides used for amplification of entire coding region**

5'-1	TTTGTTTAAACCTTGGGAAACTGG
5'-2	GTCCAGGTTTTGCTTTGATCC

FIG. 9. DNA Sequence Of Clones 4 & 7, Identical Clones Containing The Entire Open Reading Frame

```
  1  TTTGTTTAAA CCTTGGGAAA CTGGTTCAGG TCCAGGTTTT GCTTTGATCC TTTTCAAAAA
     AAACAAATTT GGAACCCTTT GACCAAGTCC AGGTCCAAAA CGAAACTAGG AAAAGTTTTT

 61  CTGGAGACAC AGAAGAGGGC TCTAGGAAAA AGTTTTGGAT GGGATTATGT GGAAACTACC
     GACCTCTGTG TCTTCTCCCG AGATCCTTTT TCAAAACCTA CCCTAATACA CCTTTGATGG

121  CTGCGATTCT CTGCTGCCAG AGCAGGCTCG GCGCTTCCAC CCCAGTGCAG CCTTCCCCTG
     GACGCTAAGA GACGACGGTC TCGTCCGAGC CGCGAAGGTG GGGTCACGTC GGAAGGGGAC

181  GCGGTGGTGA AAGAGACTCG GGAGTCGCTG CTTCCAAAGT GCCCGCCGTG AGTGAGCTCT
     CGCCACCACT TTCTCTGAGC CCTCAGCGAC GAAGGTTTCA CGGGCGGCAC TCACTCGAGA

  +2                  MetSer LeuPheGly LeuLeuLeu LeuThrSerAla LeuAlaGl
                      ]----------------------------------------------
241  CACCCCAGTC AGCCAAATGA GCCTCTTCGG GCTTCTCCTG CTGACATCTG CCCTGGCCGG
     GTGGGGTCAG TCGGTTTACT CGGAGAAGCC CGAAGAGGAC GACTGTAGAC GGGACCGGCC

  +2 yGlnArgGln GlyThrGlnAla GluSerAsn LeuSerSer LysPheGlnPhe SerSerAs
     ------------------------------------------------------------
301  CCAGAGACAG GGGACTCAGG CGGAATCCAA CCTGAGTAGT AAATTCCAGT TTTCCAGCAA
     GGTCTCTGTC CCCTGAGTCC GCCTTAGGTT GGACTCATCA TTTAAGGTCA AAAGGTCGTT

  +2 nLysGluGln AsnGlyValGln AspProGln HisGluArg IleIleThrVal SerThrAs
     ------------------------------------------------------------
361  CAAGGAACAG AACGGAGTAC AAGATCCTCA GCATGAGAGA ATTATTACTG TGTCTACTAA
     GTTCCTTGTC TTGCCTCATG TTCTAGGAGT CGTACTCTCT TAATAATGAC ACAGATGATT

  +2 nGlySerIle HisSerProArg PheProHis ThrTyrPro ArgAsnThrVal LeuValTr
     ------------------------------------------------------------
421  TGGAAGTATT CACAGCCCAA GGTTTCCTCA TACTTATCCA AGAAATACGG TCTTGGTATG
     ACCTTCATAA GTGTCGGGTT CCAAAGGAGT ATGAATAGGT TCTTTATGCC AGAACCATAC

  +2 pArgLeuVal AlaValGluGlu AsnValTrp IleGlnLeu ThrPheAspGlu ArgPheGl
     ------------------------------------------------------------
481  GAGATTAGTA GCAGTAGAGG AAAATGTATG GATACAACTT ACGTTTGATG AAAGATTTGG
     CTCTAATCAT CGTCATCTCC TTTTACATAC CTATGTTGAA TGCAAACTAC TTTCTAAACC

  +2 yLeuGluAsp ProGluAspAsp IleCysLys TyrAspPhe ValGluValGlu GluProSe
     ------------------------------------------------------------
541  GCTTGAAGAC CCAGAAGATG ACATATGCAA GTATGATTTT GTAGAAGTTG AGGAACCCAG
     CGAACTTCTG GGTCTTCTAC TGTATACGTT CATACTAAAA CATCTTCAAC TCCTTGGGTC

  +2 rAspGlyThr IleLeuGlyArg TrpCysGly SerGlyThr ValProGlyLys GlnIleSe
     ------------------------------------------------------------
601  TGATGGAACT ATATTAGGGC GCTGGTGTGG TTCTGGTACT GTACCAGGAA AACAGATTTC
     ACTACCTTGA TATAATCCCG CGACCACACC AAGACCATGA CATGGTCCTT TTGTCTAAAG

  +2 rLysGlyAsn GlnIleArgIle ArgPheVal SerAspGlu TyrPheProSer GluProGl
     ------------------------------------------------------------
661  TAAAGGAAAT CAAATTAGGA TAAGATTTGT ATCTGATGAA TATTTTCCTT CTGAACCAGG
     ATTTCCTTTA GTTTAATCCT ATTCTAAACA TAGACTACTT ATAAAAGGAA GACTTGGTCC
```

FIG. 9(CONTINUED).

```
   +2 yPheCysIle HisTyrAsnIle ValMetPro GlnPheThr GluAlaValSer ProSerVa
      ------------------------------------------------------------------
 721  GTTCTGCATC CACTACAACA TTGTCATGCC ACAATTCACA GAAGCTGTGA GTCCTTCAGT
      CAAGACGTAG GTGATGTTGT AACAGTACGG TGTTAAGTGT CTTCGACACT CAGGAAGTCA

   +2 lLeuProPro SerAlaLeuPro LeuAspLeu LeuAsnAsn AlaIleThrAla PheSerTh
      ------------------------------------------------------------------
 781  GCTACCCCCT TCAGCTTTGC CACTGGACCT GCTTAATAAT GCTATAACTG CCTTTAGTAC
      CGATGGGGGA AGTCGAAACG GTGACCTGGA CGAATTATTA CGATATTGAC GGAAATCATG

   +2 rLeuGluAsp LeuIleArgTyr LeuGluPro GluArgTrp GlnLeuAspLeu GluAspLe
      ------------------------------------------------------------------
 841  CTTGGAAGAC CTTATTCGAT ATCTTGAACC AGAGAGATGG CAGTTGGACT TAGAAGATCT
      GAACCTTCTG GAATAAGCTA TAGAACTTGG TCTCTCTACC GTCAACCTGA ATCTTCTAGA

   +2 uTyrArgPro ThrTrpGlnLeu LeuGlyLys AlaPheVal PheGlyArgLys SerArgVa
      ------------------------------------------------------------------
 901  ATATAGGCCA ACTTGGCAAC TTCTTGGCAA GGCTTTTGTT TTTGGAAGAA AATCCAGAGT
      TATATCCGGT TGAACCGTTG AAGAACCGTT CCGAAAACAA AAACCTTCTT TTAGGTCTCA

   +2 lValAspLeu AsnLeuLeuThr GluGluVal ArgLeuTyr SerCysThrPro ArgAsnPh
      ------------------------------------------------------------------
 961  GGTGGATCTG AACCTTCTAA CAGAGGAGGT AAGATTATAC AGCTGCACAC CTCGTAACTT
      CCACCTAGAC TTGGAAGATT GTCTCCTCCA TTCTAATATG TCGACGTGTG GAGCATTGAA

   +2 eSerValSer IleArgGluGlu LeuLysArg ThrAspThr IlePheTrpPro GlyCysLe
      ------------------------------------------------------------------
1021  CTCAGTGTCC ATAAGGGAAG AACTAAAGAG AACCGATACC ATTTTCTGGC CAGGTTGTCT
      GAGTCACAGG TATTCCCTTC TTGATTTCTC TTGGCTATGG TAAAAGACCG GTCCAACAGA

   +2 uLeuValLys ArgCysGlyGly AsnCysAla CysCysLeu HisAsnCysAsn GluCysGl
      ------------------------------------------------------------------
1081  CCTGGTTAAA CGCTGTGGTG GGAACTGTGC CTGTTGTCTC CACAATTGCA ATGAATGTCA
      GGACCAATTT GCGACACCAC CCTTGACACG GACAACAGAG GTGTTAACGT TACTTACAGT

   +2 nCysValPro SerLysValThr LysLysTyr HisGluVal LeuGlnLeuArg ProLysTh
      ------------------------------------------------------------------
1141  ATGTGTCCCA AGCAAAGTTA CTAAAAAATA CCACGAGGTC CTTCAGTTGA GACCAAAGAC
      TACACAGGGT TCGTTTCAAT GATTTTTTAT GGTGCTCCAG GAAGTCAACT CTGGTTTCTG

   +2 rGlyValArg GlyLeuHisLys SerLeuThr AspValAla LeuGluHisHis GluGluCy
      ------------------------------------------------------------------
1201  CGGTGTCAGG GGATTGCACA AATCACTCAC CGACGTGGCC CTGGAGCACC ATGAGGAGTG
      GCCACAGTCC CCTAACGTGT TTAGTGAGTG GCTGCACCGG GACCTCGTGG TACTCCTCAC

   +2 sAspCysVal CysArgGlySer ThrGlyGly
      ---------------------------------->
1261  TGACTGTGTG TGCAGAGGGA GCACAGGAGG ATAGCCGCAT CACCACCAGC AGCTCTTGCC
      ACTGACACAC ACGTCTCCCT CGTGTCCTCC TATCGGCGTA GTGGTGGTCG TCGAGAACGG

1321  CAGAGCTGTG CAGTGCAGTG GCTGATTCTA TTAGAGAACG TATGCGTTAT CTCCATCCTT
      GTCTCGACAC GTCACGTCAC CGACTAAGAT AATCTCTTGC ATACGCAATA GAGGTAGGAA

1381  AATCTCAGTT GTTTGCTTCA AGGACCTTTC ATCTTCAGGA TTTACAGTGC ATTCTGAAAG
      TTAGAGTCAA CAAACGAAGT TCCTGGAAAG TAGAAGTCCT AAATGTCACG TAAGACTTTC

1441  AGGAGACATC AAACAGAATT AGGAGTTGTG CAA
      TCCTCTGTAG TTTGTCTTAA TCCTCAACAC GTT
```

FIG. 10. **Predicted Full-length Polypeptide Sequence**

```
  1  MSLFGLLLLT SALAGQRQGT QAESNLSSKF QFSSNKEQYG VQDPQHERII
 51  TVSTNGSIHS PRFPHTYPRN TVLVWRLVAV EENVWIQLTF DERFGLEDPE
101  DDICKYDFVE VEEPSDGTIL GRWCGSGTVP GKQISKGNQI RIRFVSDEYF
151  PSEPGFCIHY NIVMPQFTEA VSPSVLPPSA LPLDLLNNAI TAFSTLEDLI
201  RYLEPERWQL DLEDLYRPTW QLLGKAFVFG RKSRVVDLNL LTEEVRLYSC
251  TPRNFSVSIR EELKRTDTIF WPGCLLVKRC GGNCACCLHN CNECQCVPSK
301  VTKKYHEVLQ LRPKTGVRGL HKSLTDVALE HHEECDCVCR GSTGG
```

FIG. 11. Alignment of VEGF-X with Other VEGFs

```
                        *         20         *         40         *
VEGF_HUMAN  : ----------------------------------------------------------  :   -
PLGF_HUMAN  : ----------------------------------------------------------  :   -
VEGB_HUMAN  : ----------------------------------------------------------  :   -
VEGC_HUMAN  : ----------------------------------------------------------  :   -
VEGD_HUMAN  : ----------------------------------------------------------  :   -
990126vegx  : MSLFGLLLLTSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERII          :  50

                       60         *         80         *        100
VEGF_HUMAN  : ----------------------------------------------------------  :   -
PLGF_HUMAN  : ----------------------------------------------------------  :   -
VEGB_HUMAN  : ----------------------------------------------------------  :   -
VEGC_HUMAN  : ----------------------------------------------------------  :   -
VEGD_HUMAN  : ----------------------------------------------------------  :   -
990126vegx  : TVSTNGSIHSPRFPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPE          : 100

                        *        120         *        140         *
VEGF_HUMAN  : ----------------------------------------------------------  :   -
PLGF_HUMAN  : ----------------------------------------------------------  :   -
VEGB_HUMAN  : ----------------------------------------------------------  :   -
VEGC_HUMAN  : --------------------MHLLGFFSVACSLLAAALLPGPREAPAAAA          :  30
VEGD_HUMAN  : ------------------------------------MYREWVVVNV            :  10
990126vegx  : DDICKYDFVEV--EEPSDGTILGRWCGSGTVPGKQISKGNQIRIRFVSDE          : 148

                       160         *        180         *        200
VEGF_HUMAN  : ----------------------------------------------------MN      :   2
PLGF_HUMAN  : ----------------------------------------------------MP      :   2
VEGB_HUMAN  : ----------------------------------------------------------  :   -
VEGC_HUMAN  : AFESGLDLSDAEPDAGEATAYASKDLEEQLRSVSSVDELMTVLYPEYWKM          :  80
VEGD_HUMAN  : FMMLYVQLVQGSSNEHGPVKRSSQSTLERSEQQIRAASSLEELLRITHSE          :  60
990126vegx  : YFPSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLLNNAITAFSTLED          : 198

                        *        220         *        240         *
VEGF_HUMAN  : FLLSWVHWSLALLLYLHHAKWSQAAPMAEGGGQNHHEVVKFMD-VYQRSY          :  51
PLGF_HUMAN  : VMRLFPCFLQLLAGLALPAVPPQQWALSAGNGSSEVEVVPFQE-VWGRSY          :  51
VEGB_HUMAN  : ---MSPLLRRLLLAALLQLAPAQAPVSQPDAPGHQRKVVSWID-VYTRAT          :  46
VEGC_HUMAN  : YKCQLRKGGWQHNREQANLNSRTEETIKFAAAHYNTEILKSIDNEWRKTQ          : 130
VEGD_HUMAN  : DWKLWRCRLRLKSFTSMDSRSASHRSTRFAATFYDIETLKVIDEEWQRTQ          : 110
990126vegx  : LIRYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNLLTEEVRLY          : 248

                       260         *        280         *        300
VEGF_HUMAN  : CHPIETLVDIFQEYPDEIEYIFKPSCVPLMRCGG---CCND--EGLECVP          :  96
PLGF_HUMAN  : CRALERLVDVVSEYPSEVEHMFSPSCVSLLRCTG---CCGD--ENLHCVP          :  96
VEGB_HUMAN  : CQPREVVVPLTVELMGTVAKQLVPSCVTVQRCGG---CCPD--DGLECVP          :  91
VEGC_HUMAN  : CMPREVCIDVGKEFGVATNTFFKPPCVSVYRCGG---CCNS--EGLQCMN          : 175
VEGD_HUMAN  : CSPRETCVEVASELGKSTNTFFKPPCVNVFRCGG---CCNE--ESLICMN          : 155
990126vegx  : SCTPRNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVP          : 298
```

FIG. 11 (CONTINUED).

```
                        *       320         *       340         *
VEGF_HUMAN  :  TEESNITMQIMRIKPHQG-----QHIGEMSFLQHNKCECRPKKDRARQEK  : 141
PLGF_HUMAN  :  VETANVTMQLLKIRSGDR-----PSYVELTFSQHVRCECRPLREKMKPER  : 141
VEGB_HUMAN  :  TGQHQVRMQILMIRYPS------SQLGEMSLEEHSQCECRPKKKDSAVKP  : 135
VEGC_HUMAN  :  TSTSYLSKTLFEITVPLSQG---PKPVTISFANHTSCRCMSKLDVYRQVH  : 222
VEGD_HUMAN  :  TSTSYISKQLFEISVPLTSV---PELVPVKVANHTGCKCLPTAPRHPYSI  : 202
990126vegx  :  SKVTKKYHEVLQLRPKTGVRGLHKSLTDVALEHHEECDCVCRGSTGG---  : 345

                    360         *       380         *       400
VEGF_HUMAN  :  KSVRGKGKGQKRKRKKSRYKSWSVP----------------------  : 166
PLGF_HUMAN  :  ------------------------------------------------  :   -
VEGB_HUMAN  :  DSPR--------------------------------------------  : 139
VEGC_HUMAN  :  SIIRRSLPATLPQCQAANKTCPTNYMWNNHICRCLAQEDFMFSSDAGDDS  : 272
VEGD_HUMAN  :  IRRSIQIPEEDRCSHSKKLCPIDMLWDSNKCKCVLQEENPLAGT------  : 246
990126vegx  :  ------------------------------------------------  :   -

                        *       420         *       440         *
VEGF_HUMAN  :  ------------------------------------------------  :   -
PLGF_HUMAN  :  ------------------------------------------------  :   -
VEGB_HUMAN  :  ------------------------------------------------  :   -
VEGC_HUMAN  :  TDGFHDICGPNKELDEETCQCVCRAGLRPASCGPHKELDRNSCQCVCKNK  : 322
VEGD_HUMAN  :  ------------------------------------EDHSHLQEPALCGP  : 260
990126vegx  :  ------------------------------------------------  :   -

                    460         *       480         *       500
VEGF_HUMAN  :  ---------CGPCSERRKHLFVQDPQTCKC-SCKNTDSRCKARQLELNER  : 206
PLGF_HUMAN  :  ---------CGDAVPRR------------------------------  : 149
VEGB_HUMAN  :  ----------PLCPRCTQHHQRPDPRTCRCRCRRRSFLRCQGRGLELNPD  : 179
VEGC_HUMAN  :  LFPSQCGANREFDENTCQCVCKRTCPRNQPLNPGKCACECTESPQKCLLK  : 372
VEGD_HUMAN  :  HMMFDEDRCECVCKTPCPKDLIQHPKNCSCFECKESLETCCQKHKLFHPD  : 310
990126vegx  :  ------------------------------------------------  :   -

                        *       520         *       540         *
VEGF_HUMAN  :  TCRCDKPRR---------------------------------------  : 215
PLGF_HUMAN  :  ------------------------------------------------  :   -
VEGB_HUMAN  :  TCRCRKLRR---------------------------------------  : 188
VEGC_HUMAN  :  GKKFHHQTCSCYRRPCTNRQKACEPGFSYSEEVCRCVPSYWKRPQMS---  : 419
VEGD_HUMAN  :  TCSCEDRCPFHTRPCASGKTACAKHCRFPKEKRAAQGPHSRKNP------  : 354
990126vegx  :  ------------------------------------------------  :   -
```

FIG. 12. Variant Polypeptide Sequences

```
                  *         20          *         40          *
FL_seq   : MSLFGLLLLTSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERII : 50
clone41  : MSLFGLLLLTSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERII : 50
clone20  : MSLFGLLLLTSALAGQRQGTQAESNLSSKFQFSSNKEQNGVQDPQHERII : 50

            60          *         80          *         100
FL_seq   : TVSTNGSIHSPRFPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPE : 100
clone41  : TVSTNGSIHSPRFPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPE : 100
clone20  : TVSTNGSIHSPRFPHTYPRNTVLVWRLVAVEENVWIQLTFDERFGLEDPE : 100

                  *         120         *         140         *
FL_seq   : DDICKYDFVEVEEPSDGTILGRWCGSGTVPGKQISKGNQIRIRFVSDEYF : 150
clone41  : DDICKYDFVEVEEPSDGTILGRWCGSGTVPGKQISKGNQIRIRFVSDEYF : 150
clone20  : DDICKYDFVEVEEPSDGTILGRWCGSGTVPGKQISKGNQIRIRFVSDEYF : 150

            160         *         180         *         200
FL_seq   : PSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLLNNAITAFSTLEDLI : 200
clone41  : PSEPSNRGGKIIQLHTS-------------------------------- : 167
clone20  : PSEPGFCIHYNIVMPQFTEAVSPSVLPPSALPLDLLNNAITAFSTLEDLI : 200

                  *         220         *         240         *
FL_seq   : RYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNLLTEEVRLYSC : 250
clone41  : ---------------------------------------------------- : -
clone20  : RYLEPERWQLDLEDLYRPTWQLLGKAFVFGRKSRVVDLNLLTE------- : 243

            260         *         280         *         300
FL_seq   : TPRNFSVSIREELKRTDTIFWPGCLLVKRCGGNCACCLHNCNECQCVPSK : 300
clone41  : ---------------------------------------------------- : -
clone20  : ---------------------------------------------------- : -

                  *         320         *         340
FL_seq   : VTKKYHEVLQLRPKTGVRGLHKSLTDVALEHHEECDCVCRGSTGG : 345
clone41  : ------------------------------------------- : -
clone20  : ------EVLQLRPKTGVRGLHKSLTDVALEHHEECDCVCRGSTGG : 282
```

FIG. 13. **Primers for Expression of VEGF-X**

E.coli expression of domain-

vegx-6	AATTGGATCCGAGAGTGGTGGATCTGAACC
vegx-7	AATTGGATCCGGGAAGAAAATCCAGAGTGG
vegx-8	GGTTGAATTCATTATTTTTTAGTAACTTTGCTTGGGACAC
vegX-9	AATTGAATTCATTATCCTCCTGTGCTCCCTC

Baculovirus/insect cell expression of full-length protein-

vegbac1
AATTGGATCCGGAGTCTCACCATCACCACCATCATGAATCCAACCTGAGTAGTAAATTC
C

vegbac2	AATTGAATTCGCTATCCTCCTGTGCTCCCTCTGC

FIG. 14.

>3993180H1 LUNGNON03 INCYTE
CACAAATCACTCACCGACGTGGCCCTGGAGCACCATGAGGNGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCC
GCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCAT
CCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAG
AATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCTAAAGGACAGGAGAANAGGTCTT
>3510192H1 CONCNOT01 INCYTE
TGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTT
TCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAG
GAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATTAAATAGATCACCAGCTAGTT
TCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTT
>2559870H1 ADRETUT01 INCYTE
CACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGCACCA
TGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGGGGATAGCCGCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGC
AGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCA
TCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGA
>3979767H1 LUNGTUT08 INCYTE
GGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGC
GTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAG
ACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTG
GAAAGAANATTAAATGTTGTATTAAATAGACACCAGCT
>3980011H1 LUNGTUT08 INCYTE
GGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGC
GTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACATGCATTCTGAAAGAGGAGA
CATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGG
AAAGAAAATTAAATGTTGTATTAAATAGATCACCA
>4825396H1 BLADDIT01 INCYTE
GAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTGGTGGGAACTGTGCCTGTTGTCTCCACAATT
GCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACCACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTC
AGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGG
AGGATAGCCGCATCACCACCA
>3073703H1 BONEUNT01 INCYTE
AGAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACTTCTCAGT
GTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTGGTGGGAACT
GTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACCACGAGGTCCTTCAG
TTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCA
>1302516H1 PLACNOT02 INCYTE
AGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCCTTCTGAACCTTCTAACAGAGGAGGTAAGATTATAC
AGCTGCACACCTCGTAACTTCTCAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCT
CCTGGTTAAACGCTGTGGTGGGAACTGTGCCTGTTGTCTCCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTT
ACTAAAAAATACCACGAGGTCC
>3684109H1 HEAANOT01 INCYTE
ATTTCATCTTCAGGATTTACAGTGCATTCTGAAANAGGAGAAATCAAACANAATTAGGAGTTGTGCAACAGCTCTTTTGA
GAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAANAAAATTAAATGTTGTATTAAATAGATCACCAGCTA
GTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACGGCTTAGGGTAATGTCAG
TACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTT
>4713188H1 BRAIHCT01 INCYTE
CAAAGTTACTAAAAAATACCACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCACTCACCG
ACGTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCCGCATCACCACCAGCAG
CTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGT
TTGCT
>458823H1 KERANOT01 INCYTE
ANGAGTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTT
GTTTGNTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTG
CAACAGCTCTTTTGAGAGGAGGCCTAAAGGNCAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATTAA
ATAGATC
>1303909H1 PLACNOT02 INCYTE
AGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCCTTCTGAACCTTCTAACAGAGGAGGTAAGATTATAC
AGCTGCACACCTCGTAACTTCTCAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCT
CCTGGTTAAACGCTGTGGTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAG

FIG. 14 (CONTINUED).

>2739211H1 OVARNOT09 INCYTE
GTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGA
GAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATTAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACG
TATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACGGCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAA
GTGAGCACCTGAT
>3325591H1 PTHYNOT03 INCYTE
TGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATT
AAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACG
GCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACCCTAAAGCNCC
ATGTCNNGGGCNAAAANCGAAAAAT
>3733565H1 SMCCNOS01 INCYTE
CCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGNAAGANGAGACATCAAACAG
AATTAGGNGTTGTGCAAAAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTNCAATCGTGGAAAGNAAATT
AAATGTTGTATNAAATNGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGNCNGTATTCAGTCT
TTCGGAACGGCTTAGGGTAATGTCAGTACAGGANAAAAACTGTGCAGTGAG
>3554223H1 SYNONOT01 INCYTE
ATTAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGAT
ACGGCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGGCTTAACTCTAAAG
CTCCATGTCCTGGGCCTAAAATCGTATAAAATCTGGATTTTTTTNTTTTTTTTTGCGCATATTCACATATGTAAACCAGN
ACATTCTATGTACNACAAACCTGGTTTTTAAAAAGGAAC
>4507477H1 OVARTDT01 INCYTE
GGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACGGCTTAGGGTAAT
GTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCCATGTCCTGGGCC
TAAAATCGTATAAAATCTGGA
>4163378H1 BRSTNOT32 INCYTE
AATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGNTCTGTATTTCAGTTCCTTTCGATACG
GCTTAGGGTAATGTCAGTACAGGAAAAAAGCTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCC
ATGTCCTGGGCCTAAAATCGTATA

FIG. 15.

>2054675H1 BEPINOT01 INCYTE
AAAGGAACTATGTTGCTATGAATTAAACTTGTGTCGTGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAAAATT
TCTGCCATTTAGAAGAAGAGAACTACATTCATGGTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTT
TATCGATAAGTCAGTTTATTTGTTTCATTGTGTACATTTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAA
TCTTGTTAAATATATCTATTTTTACCAAAGGTATTTAATATTCTTTTTTA
>3993180H1 LUNGNON03 INCYTE
CACAAATCACTCACCGACGTGGCCCTGGAGCACCATGAGGNGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCC
GCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCAT
CCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAG
AATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCTAAAGGACAGGAGAANAGGTCTT
>3510192H1 CONCNOT01 INCYTE
TGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTT
TCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAG
GAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATTAAATAGATCACCAGCTAGTT
TCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTT
>4164633H1 BRSTNOT32 INCYTE
CTTGTTAAATATATCTATTTTTACCAAAGGTATTTAATATTCTTTANTTATGACAACTTAGATCAACTATTTTTAGCTTG
GTAAATTTTTCTAAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATG
TATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCA
AAGACTTTTTGANAATAATTAAATTATCATATCTTCCATTCCTGTTATTGGGGGAGAAAAT
>2559870H1 ADRETUT01 INCYTE
CACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGCACCA
TGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGGGGATAGCCGCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGC
AGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCA
TCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGA
>3817470H1 BONSTUT01 INCYTE
TTAAAAAGGAACTATGTTGCTATGAATTAAACTTGTGTCATGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAA
AATTTCTGCCATTTAGAAGAAGAGAACTACATTCATGGTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTC
ACTTTATCGATAAGTCAGTTTATTTGTTTCATTGTGTACATTTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTC
TAATCTGTTAAATATATCTATTTTTACCAAAGGTATTTAATATTCTTT
>3979767H1 LUNGTUT08 INCYTE
GGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGC
GTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAG
ACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTG
GAAAGAANATTAAATGTTGTATTAAATAGACACCAGCT
>3980011H1 LUNGTUT08 INCYTE
GGAGGATAGCCGCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGC
GTTATCTCCATCCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACATGCATTCTGAAAGAGGAGA
CATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGG
AAAGAAAATTAAATGTTGTATTAAATAGATCACCA
>4825396H1 BLADDIT01 INCYTE
GAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTGGTGGGAACTGTGCCTGTTGTCTCCACAATT
GCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACCACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTC
AGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGG
AGGATAGCCGCATCACCACCA
>3073703H1 BONEUNT01 INCYTE
AGAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACTTCTCAGT
GTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTGGTGGGAACT
GTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACCACGAGGTCCTTCAG
TTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCA
>862169H1 BRAITUT03 INCYTE
AGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTGCA
TTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAAGACTTTTTGAAAATAATTAAATTATCATATCTTCCATTC
CTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGTAGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTT
GGGGAAATCTGAGCCTAGC
>4201385H1 BRAITUT29 INCYTE
TTTTTAAAAAGGAACTATGTTGCTATGAATTAAACTTGTGTCGTGCTGATAGGACAGACTGGATTTTTCATATTTCTTAT
TAAAATTTCTGCCATTTAGAAGAAGAGAACTACATTCATGGTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTATCT
TCACTTTATCGATAAGTCAGTTTATTTGTTTCATTGTGTACATTTTTATATTCTCCTTTGACATATAACTGTTGGCTTTT

FIG. 15 (CONTINUED 1).

CTAATCTGTTAAATATATCTATTTTTACCAAAGGTATTTAATAT
>1302516H1 PLACNOT02 INCYTE
AGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCCTTCTGAACCTTCTAACAGAGGAGGTAAGATTATAC
AGCTGCACACCTCGTAACTTCTCAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCT
CCTGGTTAAACGCTGTGGTGGGAACTGTGCCTGTTGTCTCCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTT
ACTAAAAAATACCACGAGGTCC
>3684109H1 HEAANOT01 INCYTE
ATTTCATCTTCAGGATTTACAGTGCATTCTGAAANAGGAGAAATCAAACANAATTAGGAGTTGTGCAACAGCTCTTTTGA
GAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAANAAAATTAAATGTTGTATTAAATAGATCACCAGCTA
GTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACGGCTTAGGGTAATGTCAG
TACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTT
>2549720H1 LUNGTUT06 INCYTE
TTAGCTTGGNAAATTTTTCTAAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAA
AATACATGTATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGT
AAGTTGCAAAGACTTTTTGAAAATAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACT
TATGANAGTAG
>877279H1 LUNGAST01 INCYTE
CTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCTAAACACAATTGTTATAGCCAGAGGAACAAA
GATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTGCAT
TTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAAGGCTTTTTGAAAATAATTAAATTATCATATCTTCCATTCC
TGTTATTGGNGG
>4713188H1 BRAIHCT01 INCYTE
CAAAGTTACTAAAAAATACCACGAGGTCCTTCAGTTGAGACCAAAGACCGGTGTCAGGGGATTGCACAAATCACTCACCG
ACGTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGGATAGCCGCATCACCACCAGCAG
CTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTTGT
TTGCT
>2171082H1 ENDCNOT03 INCYTE
AGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTGTTTCATTGTGTACATTTTTA
TATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTAAATATATCTATTTTTACCAAAGGTATTTAATATT
CTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCTAAACACAATTGTTATAGCCAGAGGAACAAA
GATGA
>875860H1 LUNGAST01 INCYTE
CTGGATTTTTCATATTTCTTATTAAAATTTCTGCCATTTAGAAGAAGAGAACTACATTCATGGTTTGGAAGAGATAAACC
TGAAAAGAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTGTTTCATTGTGTACATTTTTATATTCTCCT
TTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTAAATATATCTATTTTTACCAAAGGTATTTAATATTCTTTTTTAT
GAC
>706168H1 SYNORAT04 INCYTE
GCTCATATTCACATATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGANCTATGTTGCTATGAAT
TAAACTTGTGTCGTGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAAAATTTCTGCCATTTAGAAGAAGAGAAC
TACATTCATGGTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCANTTTATCGATAAGTCAGTTTATTTGT
TTCA
>458823H1 KERANOT01 INCYTE
ANGAGTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCATCCTTAATCTCAGTT
GTTTGNTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTG
CAACAGCTCTTTTGAGAGGAGGCCTAAAGGNCAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATTAA
ATAGATC
>538436H1 LNODNOT02 INCYTE
AAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTG
CATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAAGACTTTTTGAAAATAATTAAATTATCATATCTTCCAT
TCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGTAGACATTCAGATCCAGCCATTACTAACCTAT
>1303909H1 PLACNOT02 INCYTE
AGGAAATCAAATTAGGATAAGATTTGTATCTGATGAATATTTTCCTTCTGAACCTTCTAACAGAGGAGGTAAGATTATAC
AGCTGCACACCTCGTAACTTCTCAGTGTCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCT
CCTGGTTAAACGCTGTGGTGGGAACTGTGCCTGTTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAG
>2739211H1 OVARNOT09 INCYTE
GTGCATTCTGAAAGAGGAGACATCAAACAGAATTAGGAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGA
GAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATTAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACG
TATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACGGCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAA
GTGAGCACCTGAT

FIG. 15(CONTINUED 2).

>2550343H1 LUNGTUT06 INCYTE
TGTACATTTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCNAATCTTGTTAAATATATCTATTTTTACCAAAG
GTATTTAATATTCTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCTAAACACAATTGTTATAGC
CAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCATTCTCGTATGGTGCTA
>5321148H1 FIBPFEN06 INCYTE
CACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGNCAAAAATACATGTATTTCATTCTCGTA
TGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAAGACTTTTTGAAAA
TAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGTAAATTCAGATCCAC
CATTACTAAC
>879495H1 THYRNOT02 INCYTE
ATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAA
AGACTTTTTGAAAATAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGT
AGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCT
TGAAAAA
>3325591H1 PTHYNOT03 INCYTE
TGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATT
AAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACG
GCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACCCTAAAGCNCC
ATGTCNNGGGCNAAAANCGAAAAAT
>543890H1 OVARNOT02 INCYTE
TTTCTAAACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCA
TTCTCGTATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGNATAGAATTGGTAAGTTGCAAAGNCTT
TTTGAAAATAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGGATGGAAAATAAAAAGCAACTTATGGAAAGTAGG
ACATTCAGATC
>3733565H1 SMCCNOS01 INCYTE
CCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGNAAGANGAGACATCAAACAG
AATTAGGNGTTGTGCAAAAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTNCAATCGTGGAAAGNAAATT
AAATGTTGTATNAAATNGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGNCNGTATTCAGTCT
TTCGGAACGGCTTAGGGTAATGTCAGTACAGGANAAAAACTGTGCAGTGAG
>4641939H1 PROSTMT03 INCYTE
GTACTACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAAACTTGTGTCCATGCTGATAGGACAGACTGGAT
TTTNCATATTTCTTATTAAAATTTCTGCCATTTAGAAGAAGAGAACTACATTCATGGTTTGGNAGAGATAAACCTGAAAA
GAAGAGTGGCCTTATCTTCACTTTATCGATAAGTCAGTTTATTTGTTTCATGTGTACATTTTTATATTCTCCTTTGACAT
ATAACGTGGCTTT
>2007780H1 TESTNOT03 INCYTE
TTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTAAATATATCTATTTTTACCAAAGGTATTTAAT
ATTCTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCTAAACACAATTGTTATAGCCAGAGGAAC
AAAGATGATATAAAATATTGTTGCTCTGANAAAAATACATGTAT
>3085331H1 HEAONOT03 INCYTE
GCTCATATTCACATATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACTATTTGCTATGAATT
AAACTTGTGTCGTGCTGATAGGACAGACTGGNTTTTTCATATTTCTTATTANAATTTCTGCCATTAGAAGAAGAGAACTA
CATTCATGGTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTATTTCACTTTATCGATAAGTCAGT
>3414043H1 PTHYNOT04 INCYTE
GCTCATATTCACATATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAAT
TAAACTTGTGTCGTGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAAAATTTCTGCCATTTAGAAGAAGAGAAC
TACATTCATGGTTTGGAAGAGATAAACCTGAAA
>3705963H1 PENCNOT07 INCYTE
ANACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCCATGTCCTGGGCCTAAAATCGTATAAAA
TCTGGAnnnnnnnnnnnnnnnnnnnGCTCATATTCACATATGTAAACCAGAACATTCTATGTACTACAAACCTGGTTTTTA
AAAAGGAACTATGTTGCTATGAATTAAACTTGTGTCGTGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAAAAT
TTCTGCCATTAGAAGAAGAGAACTACNTTCANGGTTTGGAAGAGATAACCCTGAAAAGANGGG
>5137051H1 OVARDIT04 INCYTE
AAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAAGACTNTTTGAAAATAATTAAATTATCATATCTTCCATTCCTGT
TATTGGAGATGAANATAAAAAGCAACTTATGAAAGTAGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTTGGGG
AAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCGTGCTGTNTGTCA
GTAGGAACACATCCTATTTATTGTGATGNTGTGGTTTATTAT
>3554223H1 SYNONOT01 INCYTE
ATTAAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGAT
ACGGCTTAGGGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGGCTTAACTCTAAAG

FIG. 15 (CONTINUED 3).

CTCCATGTCCTGGGCCTAAAATCGTATAAAATCTGGATTTTTTTNTTTTTTTTTGCGCATATTCACATATGTAAACCAGN
ACATTCTATGTACNACAAACCTGGTTTTTAAAAAGGAAC
>4507477H1 OVARTDT01 INCYTE
GGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACGGCTTAGGGTAAT
GTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCCATGTCCTGGGCC
TAAAATCGTATAAAATCTGGA
>1955646H1 CONNNOT01 INCYTE
TGGTAAGTTGCAAAGACTTTTTGAAAATAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGC
AACTTATGAAAGTAGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAA
ACATAAAGCACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCGTGCTGTGCTGTGCAGTAGGGAACACATCCTATTTA
TTGTGATGTTGTGGTTTATATCCTAAACC
>4163378H1 BRSTNOT32 INCYTE
AATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGNTCTGTATTTCAGTTCCTTTCGATACG
GCTTAGGGTAATGTCAGTACAGGAAAAAAGCTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCC
ATGTCCTGGGCCTAAAATCGTATA
>5095141H1 EPIMNON05 INCYTE
AGATAAACCTGAAAAGAAGAGTGGCCTTATNTTCACTTTATCGATAAGTCAGNTTATTTGTTTCATTGTGTACATTTNNA
TATTCTCCTTTTGACATTATAACTGNTGGCTTTTCTAANCNTGTTAAATATATCTATTTTTACCAAAGGTATTTAATATT
CTTT
>943826H1 ADRENOT03 INCYTE
TATGGTGCTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAAGACTTTTTGAA
AATAATTAAATTATCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATG
>3451273H1 UTRSNON03 INCYTE
TTTTTTNTTTTGCTCATATTCACATATGTAAACCNGAACATTCTATGTACNACAAACCTGGTTTTTAAAAAGGAACTATG
TTGCTATGAATTAAACTTGTGTCGTGCTGATAGGACAGACTGGATTTTTCANATTTCTTANTAANNTTTCTGCCATTTAG
AAGA
>1402278H1 LATRTUT02 INCYTE
GTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCCATGTCCTGGGCCTAAA
ATCGTATAAAATCTGGAnnnnnnnnnnnnnnnnnnnGCTCATATTCACATATGTAAACCAGAACATTCTATGTACTACAAA
CCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAAACTTGTGTCGTGCTGATAGGACAGACTGGATTTTTCATATTT
CTTA
>4361191H1 SKIRNOT01 INCYTE
GCAAAGACTTTTTGANAATNATTAANTTATCATATCTTCCATTCCTGTTATNGGAGATGANAATAAAAAGCAACTTATGA
AAGTAGACATTCAGATCCAGCCATTACTAACCTATTCCTTTTTTGGGGAAATCTGAGCCTAGCNCAGAAAAACATAAAGC
ACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCGTGCTGTGCTGTGCAGTAGGAACACATCCNATTTATTGTGNTGTN
GNGGTTTTATGATC
>1307017H1 PLACNOT02 INCYTE
TGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGCTTAACTCTAAAGCTCCATGTCCTGGGC
CTAAAATCGTATAAAATCTGGAnnnnnnnnnnnnnnnnnnnGCTCATATTCACATATGTAAACCAGAACATTCTATGTACT
ACAAACCTGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAAACTTGTGTCATGCTGATAGGACAGACTGGATTTTTCA
TAT
>5032225H1 HEARFET03 INCYTE
AATTATCATATCTTCCATTCCTGTTATTGGAGATGNAAATAAAAAGCAACTTATGAAAGTAGACATTCAGATCCAGCCAT
TACTAACCTATTCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCTTGAAAAAGACTGTCAGCTTC
CTGATAAAGCGTGCTGTGCTGTGCAGTAGGAACACATCCTATTTATTGTGATGTTGTGGTTTTATTATCTTAAACTCGTT
CCAT
>3732621H1 SMCCNOS01 INCYTE
ANAGATGATATAAAANATTGTTGCTCTGACAANNATACATGTATTTCATTCTCGTATGGTGCTAGAGTTAGATTAATCTG
CNTTTTAAAAAACTGANTTGGAATAGANTTGGTAAGTTGCAAAGNCNTTTGAAAATNATTAAGTTATCAGAT
>3530274H1 BLADNOT09 INCYTE
TTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGTAGACATTCAGATCCAGCCATTACTAACCTATT
CCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCG
TGCTGTGCTGTGCAGTAGGAACACATCCTATTTATTGTGATGTTGTGGTTTTATTATCTAAACTCTGTTCCATACACTTG
TATAAATACATGGATATTTTTATGTACAGAAGTATGTCTCTTAACCAGTTCA
>3530249H1 BLADNOT09 INCYTE
CTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGANAGTAGACATTCAGATCCAGCCATTACTAACCTAT
TCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGC
GTGCTGTGCTGTGCAGTAGGAACACATCCTATTTATTGTGATGTTGTGGTTTTATTATCTTAAACTCTGTTCCATACACT
TGTATAAATACATGGATATTTTTATGTACAGAAGTATGTCTCTTAACCAGTTCACTTATTGTACCTGG

FIG. 16.

VEGFE1	AAAATGTATGGATACAACTTAC	22
VEGFE2	GTTTGATGAAAGATTTGGGCTTG	23
VEGFE3	TTTCTAAAGGAAATCAAATTAG	22
VEGFE4	GATAAGATTTGTATCTGATG	20
VEGFE5	GATGTCTCCTCTTTCAG	17
VEGFE6	GCACAACTCCTAATTCTG	18
VEGFE7	AGCACCTGATTCCGTTGC	19
VEGFE8	TAGTACATAGAATGTTCTGG	20
VEGFE9	AAGAGACATACTTCTGTAC	19
VEGFE10	CCAGGTACAATAAGTGAACTG	21

FIG. 17.

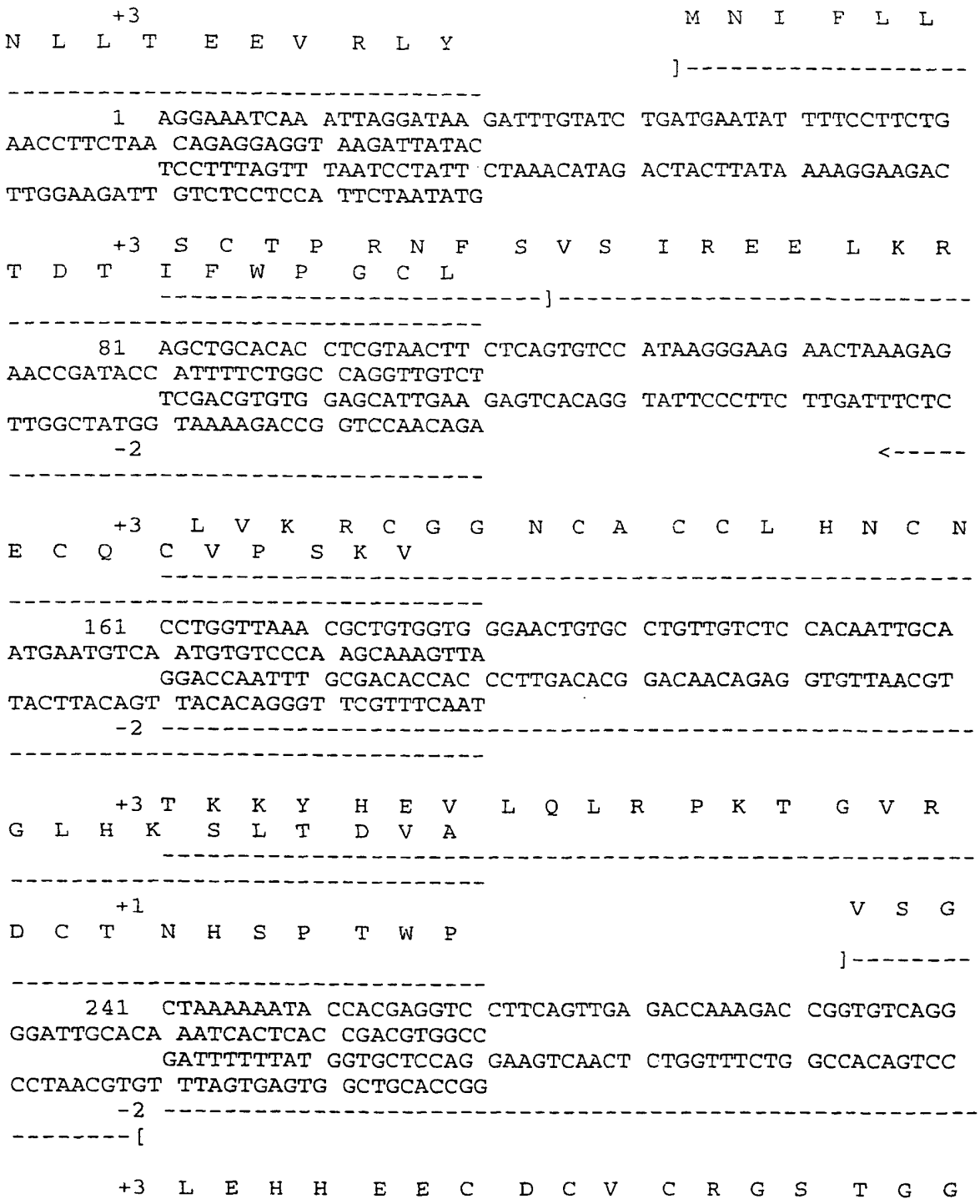

+3
M N I F L L
N L L T E E V R L Y
1 AGGAAATCAA ATTAGGATAA GATTTGTATC TGATGAATAT TTTCCTTCTG
AACCTTCTAA CAGAGGAGGT AAGATTATAC
TCCTTTAGTT TAATCCTATT CTAAACATAG ACTACTTATA AAAGGAAGAC
TTGGAAGATT GTCTCCTCCA TTCTAATATG
+3 S C T P R N F S V S I R E E L K R
T D T I F W P G C L
81 AGCTGCACAC CTCGTAACTT CTCAGTGTCC ATAAGGGAAG AACTAAAGAG
AACCGATACC ATTTTCTGGC CAGGTTGTCT
TCGACGTGTG GAGCATTGAA GAGTCACAGG TATTCCCTTC TTGATTTCTC
TTGGCTATGG TAAAAGACCG GTCCAACAGA
-2
+3 L V K R C G G N C A C C L H N C N
E C Q C V P S K V
161 CCTGGTTAAA CGCTGTGGTG GGAACTGTGC CTGTTGTCTC CACAATTGCA
ATGAATGTCA ATGTGTCCCA AGCAAAGTTA
GGACCAATTT GCGACACCAC CCTTGACACG GACAACAGAG GTGTTAACGT
TACTTACAGT TACACAGGGT TCGTTTCAAT
-2
+3 T K K Y H E V L Q L R P K T G V R
G L H K S L T D V A
+1
V S G
D C T N H S P T W P
241 CTAAAAAATA CCACGAGGTC CTTCAGTTGA GACCAAAGAC CGGTGTCAGG
GGATTGCACA AATCACTCAC CGACGTGGCC
GATTTTTTAT GGTGCTCCAG GAAGTCAACT CTGGTTTCTG GCCACAGTCC
CCTAACGTGT TTAGTGAGTG GCTGCACCGG
-2
+3 L E H H E E C D C V C R G S T G G

FIG. 17 (CONTINUED).

```
          ----------------------------------------------------
>
       +2                                    V   Q  R  E  H  R  R
I  A  A  S   P  P  A   A  L  A
                                           ]----------------------
--------------------------------
       +1   W  S  T   M  R  S  V   T  V  C   A  E  G   A  Q  E  D
S  R  I   T  T  S   S  S  C
          ----------------------------------------------------
--------------------------------
     321  CTGGAGCACC ATGAGGAGTG TGACTGTGTG TGCAGAGGGA GCACAGGAGG
ATAGCCGCAT CACCACCAGC AGCTCTTGCC
          GACCTCGTGG TACTCCTCAC ACTGACACAC ACGTCTCCCT CGTGTCCTCC
TATCGGCGTA GTGGTGGTCG TCGAGAACGG

       +2  Q  S  C  A   V  Q  W   L  I  L   L  E  N  V   C  V  I
S  I  L   N  L  S  C   L  L  Q
          ----------------------------------------------------
--------------------------------
       +1 P  E  L  C   S  A  V   A  D  S  I   R  E  R   M  R  Y
L  H  P
          ----------------------------------------------------
-------->
     401  CAGAGCTGTG CAGTGCAGTG GCTGATTCTA TTAGAGAACG TATGCGTTAT
CTCCATCCTT AATCTCAGTT GTTTGCTTCA
          GTCTCGACAC GTCACGTCAC CGACTAAGAT AATCTCTTGC ATACGCAATA
GAGGTAGGAA TTAGAGTCAA CAAACGAAGT

       +2   G  P  F   I  F  R  I   Y  S  A   F
          ----------------------------------->
     481  AGGACCTTTC ATCTTCAGGA TTTACAGTGC ATTCTGAAAG AGGAGACATC
AAACAGAATT AGGAGTTGTG CAACAGCTCT
          TCCTGGAAAG TAGAAGTCCT AAATGTCACG TAAGACTTTC TCCTCTGTAG
TTTGTCTTAA TCCTCAACAC GTTGTCGAGA

     561  TTTGAGAGGA GGCCTAAAGG ACAGGAGAAA AGGTCTTCAA TCGTGGAAAG
AAAATTAAAT GTTGTATTAA ATAGATCACC
          AAACTCTCCT CCGGATTTCC TGTCCTCTTT TCCAGAAGTT AGCACCTTTC
TTTTAATTTA CAACATAATT TATCTAGTGG

     641  AGCTAGTTTC AGAGTTACCA TGTACGTATT CCACTAGCTG GGTTCTGTAT
TTCAGTTCTT TCGATACGGC TTAGGGTAAT
          TCGATCAAAG TCTCAATGGT ACATGCATAA GGTGATCGAC CCAAGACATA
AAGTCAAGAA AGCTATGCCG AATCCCATTA

     721  GTCAGTACAG GAAAAAAACT GTGCAAGTGA GCACCTGATT CCGTTGCCTT
GGCTTAACTC TAAAGCTCCA TGTCCTGGGC
          CAGTCATGTC CTTTTTTTGA CACGTTCACT CGTGGACTAA GGCAACGGAA
CCGAATTGAG ATTTCGAGGT ACAGGACCCG

     801  CTAAAATCGT ATAAAATCTG GA
          GATTTTAGCA TATTTTAGAC CT
```

FIG. 18.

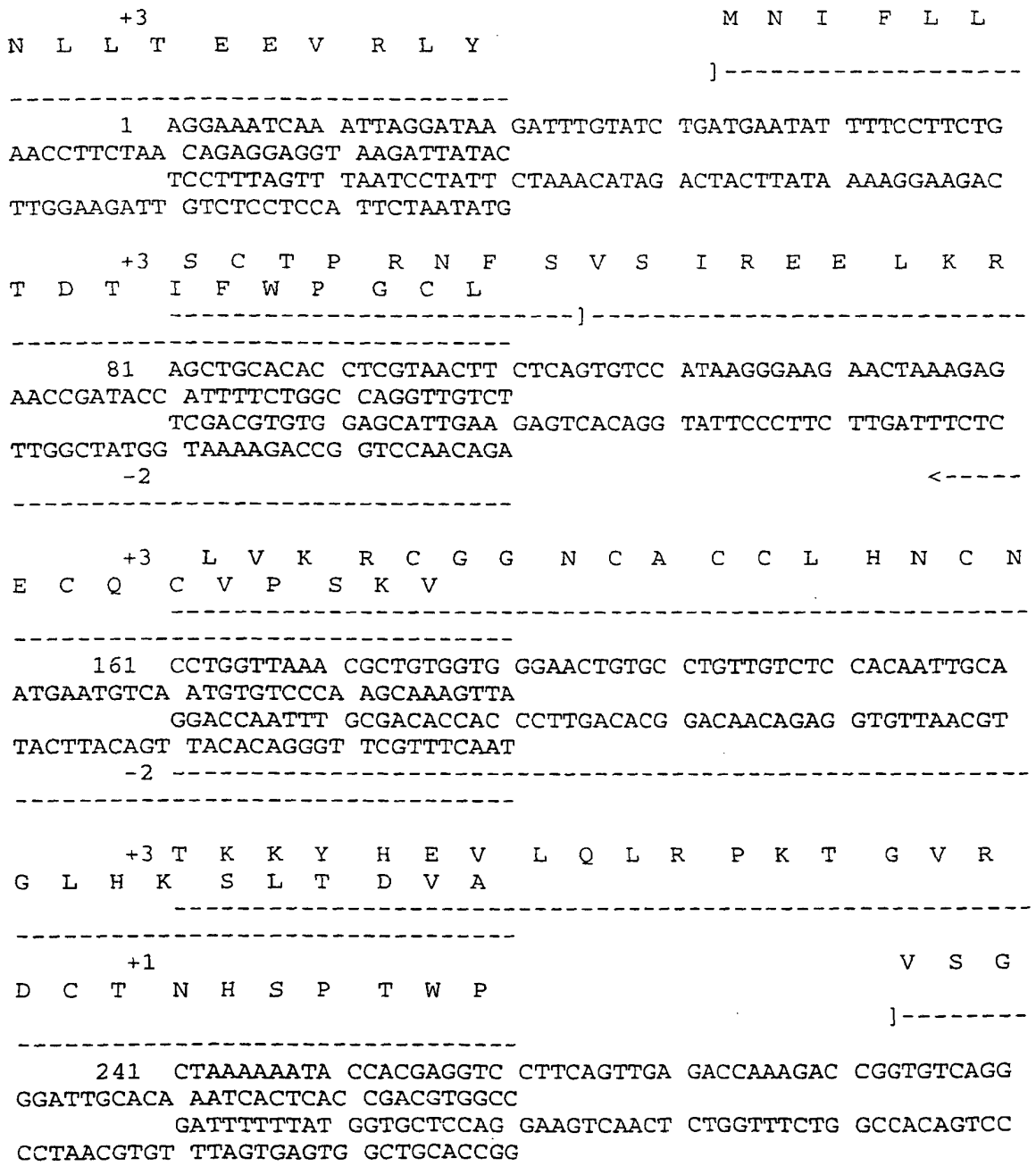

+3 M N I F L L
N L L T E E V R L Y
1 AGGAAATCAA ATTAGGATAA GATTTGTATC TGATGAATAT TTTCCTTCTG
AACCTTCTAA CAGAGGAGGT AAGATTATAC
TCCTTTAGTT TAATCCTATT CTAAACATAG ACTACTTATA AAAGGAAGAC
TTGGAAGATT GTCTCCTCCA TTCTAATATG
+3 S C T P R N F S V S I R E E L K R
T D T I F W P G C L
81 AGCTGCACAC CTCGTAACTT CTCAGTGTCC ATAAGGGAAG AACTAAAGAG
AACCGATACC ATTTTCTGGC CAGGTTGTCT
TCGACGTGTG GAGCATTGAA GAGTCACAGG TATTCCCTTC TTGATTTCTC
TTGGCTATGG TAAAAGACCG GTCCAACAGA
-2
+3 L V K R C G G N C A C C L H N C N
E C Q C V P S K V
161 CCTGGTTAAA CGCTGTGGTG GGAACTGTGC CTGTTGTCTC CACAATTGCA
ATGAATGTCA ATGTGTCCCA AGCAAAGTTA
GGACCAATTT GCGACACCAC CCTTGACACG GACAACAGAG GTGTTAACGT
TACTTACAGT TACACAGGGT TCGTTTCAAT
-2
+3 T K K Y H E V L Q L R P K T G V R
G L H K S L T D V A
+1 V S G
D C T N H S P T W P
241 CTAAAAAATA CCACGAGGTC CTTCAGTTGA GACCAAAGAC CGGTGTCAGG
GGATTGCACA AATCACTCAC CGACGTGGCC
GATTTTTTAT GGTGCTCCAG GAAGTCAACT CTGGTTTCTG GCCACAGTCC
CCTAACGTGT TTAGTGAGTG GCTGCACCGG

FIG. 1B (CONTINUED 1).

```
      -2 ----------------------------------------------------
--------[

      +3  L  E  H  H   E  E  C   D  C  V   C  R  G  S   T  G  G
          ----------------------------------------------------
>
      +2                                    V  Q  R  E   H  R  R
I  A  A  S   P  P  A   A  L  A
                                           ]----------------------
---------------------------------
      +1  W  S  T  M  R  S  V   T  V  C   A  E  G  A  Q  E  D
S  R  I  T  T  S  S  S  C
          ----------------------------------------------------
---------------------------------
    321  CTGGAGCACC ATGAGGAGTG TGACTGTGTG TGCAGAGGGA GCACAGGAGG
ATAGCCGCAT CACCACCAGC AGCTCTTGCC
         GACCTCGTGG TACTCCTCAC ACTGACACAC ACGTCTCCCT CGTGTCCTCC
TATCGGCGTA GTGGTGGTCG TCGAGAACGG

      +2  Q  S  C  A   V  Q  W   L  I  L   L  E  N  V   C  V  I
S  I  L   N  L  S  C   L  L  Q
          ----------------------------------------------------
---------------------------------
      +1 P  E  L  C   S  A  V   A  D  S  I   R  E  R   M  R  Y
L  H  P
          ----------------------------------------------------
-------->
    401  CAGAGCTGTG CAGTGCAGTG GCTGATTCTA TTAGAGAACG TATGCGTTAT
CTCCATCCTT AATCTCAGTT GTTTGCTTCA
         GTCTCGACAC GTCACGTCAC CGACTAAGAT AATCTCTTGC ATACGCAATA
GAGGTAGGAA TTAGAGTCAA CAAACGAAGT

      +2   G  P  F   I  F  R  I   Y  S  A   F
          ----------------------------------->
    481  AGGACCTTTC ATCTTCAGGA TTTACAGTGC ATTCTGAAAG AGGAGACATC
AAACAGAATT AGGAGTTGTG CAACAGCTCT
         TCCTGGAAAG TAGAAGTCCT AAATGTCACG TAAGACTTTC TCCTCTGTAG
TTTGTCTTAA TCCTCAACAC GTTGTCGAGA

    561  TTTGAGAGGA GGCCTAAAGG ACAGGAGAAA AGGTCTTCAA TCGTGGAAAG
AAAATTAAAT GTTGTATTAA ATAGATCACC
         AAACTCTCCT CCGGATTTCC TGTCCTCTTT TCCAGAAGTT AGCACCTTTC
TTTTAATTTA CAACATAATT TATCTAGTGG

    641  AGCTAGTTTC AGAGTTACCA TGTACGTATT CCACTAGCTG GGTTCTGTAT
TTCAGTTCTT TCGATACGGC TTAGGGTAAT
         TCGATCAAAG TCTCAATGGT ACATGCATAA GGTGATCGAC CCAAGACATA
AAGTCAAGAA AGCTATGCCG AATCCCATTA

    721  GTCAGTACAG GAAAAAAACT GTGCAAGTGA GCACCTGATT CCGTTGCCTT
GGCTTAACTC TAAAGCTCCA TGTCCTGGGC
```

FIG. 18 (CONTINUED 2).

```
          CAGTCATGTC CTTTTTTTGA CACGTTCACT CGTGGACTAA GGCAACGGAA
CCGAATTGAG ATTTCGAGGT ACAGGACCCG

     801  CTAAAATCGT ATAAAATCTG GATTTTTTTN TTTTTTTTTG CGCATATTCA
CATATGTAAA CCAGAACATT CTATGTACTA
          GATTTTAGCA TATTTTAGAC CTAAAAAAAN AAAAAAAAAC GCGTATAAGT
GTATACATTT GGTCTTGTAA GATACATGAT

     881  CAAACCTGGT TTTTAAAAAG GAACTATGTT GCTATGAATT AAACTTGTGT
CGTGCTGATA GGACAGACTG GATTTTTCAT
          GTTTGGACCA AAAATTTTTC CTTGATACAA CGATACTTAA TTTGAACACA
GCACGACTAT CCTGTCTGAC CTAAAAAGTA
       -3                  <--------------------------------------
--------------------------------

     961  ATTTCTTATT AAAATTTCTG CCATTTAGAA GAAGAGAACT ACATTCATGG
TTTGGAAGAG ATAAACCTGA AAAGAAGAGT
          TAAAGAATAA TTTTAAAGAC GGTAAATCTT CTTCTCTTGA TGTAAGTACC
AAACCTTCTC TATTTGGACT TTTCTTCTCA
       -3 ----------------------------------------------------------
--------------------------------

    1041  GGCCTTATCT TCACTTTATC GATAAGTCAG TTTATTTGTT TCATTGTGTA
CATTTTTATA TTCTCCTTTT GACATTATAA
          CCGGAATAGA AGTGAAATAG CTATTCAGTC AAATAAACAA AGTAACACAT
GTAAAAATAT AAGAGGAAAA CTGTAATATT
       -3 --------------[

    1121  CTGTTGGCTT TTCTAATCTT GTTAAATATA TCTATTTTTA CCAAAGGTAT
TTAATATTCT TTTTTATGAC AACTTAGATC
          GACAACCGAA AAGATTAGAA CAATTTATAT AGATAAAAAT GGTTTCCATA
AATTATAAGA AAAAATACTG TTGAATCTAG

    1201  AACTATTTTT AGCTTGGTAA ATTTTTCTAA ACACAATTGT TATAGCCAGA
GGAACAAAGA TGATATAAAA TATTGTTGCT
          TTGATAAAAA TCGAACCATT TAAAAAGATT TGTGTTAACA ATATCGGTCT
CCTTGTTTCT ACTATATTTT ATAACAACGA

    1281  CTGACAAAAA TACATGTATT TCATTCTCGT ATGGTGCTAG AGTTAGATTA
ATCTGCATTT TAAAAAACTG AATTGGAATA
          GACTGTTTTT ATGTACATAA AGTAAGAGCA TACCACGATC TCAATCTAAT
TAGACGTAAA ATTTTTTGAC TTAACCTTAT

    1361  GAATTGGTAA GTTGCAAAGA CTTTTTGAAA ATAATTAAAT TATCATATCT
TCCATTCCTG TTATTGGAGA TGAAAATAAA
          CTTAACCATT CAACGTTTCT GAAAAACTTT TATTAATTTA ATAGTATAGA
AGGTAAGGAC AATAACCTCT ACTTTTTATTT

    1441  AAGCAACTTA TGAAAGTAGA CATTCAGATC CAGCCATTAC TAACCTATTC
CTTTTTTGGG GAAATCTGAG CCTAGCTCAG
          TTCGTTGAAT ACTTTCATCT GTAAGTCTAG GTCGGTAATG ATTGGATAAG
GAAAAAACCC CTTTAGACTC GGATCGAGTC
```

FIG. 18 (CONTINUED 3).

```
1521  AAAAACATAA AGCACCTTGA AAAAGACTTG GCAGCTTCCT GATAAAGCGT
GCTGTGCTGT GCAGTAGGAA CACATCCTAT
      TTTTTGTATT TCGTGGAACT TTTTCTGAAC CGTCGAAGGA CTATTTCGCA
CGACACGACA CGTCATCCTT GTGTAGGATA

1601  TTATTGTGAT GTTGTGGTTT TATTATCTTA AACTCTGTTC CATACACTTG
TATAAATACA TGGATATTTT TATGTACAGA
      AATAACACTA CAACACCAAA ATAATAGAAT TTGAGACAAG GTATGTGAAC
ATATTTATGT ACCTATAAAA ATACATGTCT

1681  AGTATGTCTC TTAACCAGTT CACTTATTGT ACCTGG
      TCATACAGAG AATTGGTCAA GTGAATAACA TGGACC
```

FIG. 19. DNA and polypeptide sequence used for mammalian cell expression

```
 +1          m  s  l  f  g  l  l  l  l  t  s  a  l  a  g  q  r
  1  GGATCCAAAA TGAGCCTCTT CGGGCTTCTC CTGCTGACAT CTGCCCTGGC CGGCCAGAGA

 +1  q  g  t  q  a  E  S  N  L  S  S  K  F  Q  F  S  S  N  K  E
 61  CAGGGGACTC AGGCGGAATC CAACCTGAGT AGTAAATTCC AGTTTTCCAG CAACAAGGAA

 +1  Q  N  G  V  Q  D  P  Q  H  E  R  I  I  T  V  S  T  N  G  S
121  CAGAACGGAG TACAAGATCC TCAGCATGAG AGAATTATTA CTGTGTCTAC TAATGGAAGT

 +1  I  H  S  P  R  F  P  H  T  Y  P  R  N  T  V  L  V  W  R  L
181  ATTCACAGCC CAAGGTTTCC TCATACTTAT CCAAGAAATA CGGTCTTGGT ATGGAGATTA

 +1  V  A  V  E  E  N  V  W  I  Q  L  T  F  D  E  R  F  G  L  E
241  GTAGCAGTAG AGGAAAATGT ATGGATACAA CTTACGTTTG ATGAAAGATT TGGGCTTGAA

 +1  D  P  E  D  D  I  C  K  Y  D  F  V  E  V  E  E  P  S  D  G
301  GACCCAGAAG ATGACATATG CAAGTATGAT TTTGTAGAAG TTGAGGAACC CAGTGATGGA

 +1  T  I  L  G  R  W  C  G  S  G  T  V  P  G  K  Q  I  S  K  G
361  ACTATATTAG GGCGCTGGTG TGGTTCTGGT ACTGTACCAG GAAAACAGAT TTCTAAAGGA

 +1  N  Q  I  R  I  R  F  V  S  D  E  Y  F  P  S  E  P  G  F  C
421  AATCAAATTA GGATAAGATT TGTATCTGAT GAATATTTTC CTTCTGAACC AGGGTTCTGC

 +1  I  H  Y  N  I  V  M  P  Q  F  T  E  A  V  S  P  S  V  L  P
481  ATCCACTACA ACATTGTCAT GCCACAATTC ACAGAAGCTG TGAGTCCTTC AGTGCTACCC

 +1  P  S  A  L  P  L  D  L  L  N  N  A  I  T  A  F  S  T  L  E
541  CCTTCAGCTT TGCCACTGGA CCTGCTTAAT AATGCTATAA CTGCCTTTAG TACCTTGGAA

 +1  D  L  I  R  Y  L  E  P  E  R  W  Q  L  D  L  E  D  L  Y  R
601  GACCTTATTC GATATCTTGA ACCAGAGAGA TGGCAGTTGG ACTTAGAAGA TCTATATAGG

 +1  P  T  W  Q  L  L  G  K  A  F  V  F  G  R  K  S  R  V  V  D
661  CCAACTTGGC AACTTCTTGG CAAGGCTTTT GTTTTTGGAA GAAAATCCAG AGTGGTGGAT

 +1  L  N  L  L  T  E  E  V  R  L  Y  S  C  T  P  R  N  F  S  V
721  CTGAACCTTC TAACAGAGGA GGTAAGATTA TACAGCTGCA CACCTCGTAA CTTCTCAGTG

 +1  S  I  R  E  E  L  K  R  T  D  T  I  F  W  P  G  C  L  L  V
781  TCCATAAGGG AAGAACTAAA GAGAACCGAT ACCATTTTCT GGCCAGGTTG TCTCCTGGTT

 +1  K  R  C  G  G  N  C  A  C  C  L  H  N  C  N  E  C  Q  C  V
841  AAACGCTGTG GTGGGAACTG TGCCTGTTGT CTCCACAATT GCAATGAATG TCAATGTGTC

 +1  P  S  K  V  T  K  K  Y  H  E  V  L  Q  L  R  P  K  T  G  V
901  CCAAGCAAAG TTACTAAAAA ATACCACGAG GTCCTTCAGT TGAGACCAAA GACCGGTGTC

 +1  R  G  L  H  K  S  L  T  D  V  A  L  E  H  H  E  E  C  D  C
961  AGGGGATTGC ACAAATCACT CACCGACGTG GCCCTGGAGC ACCATGAGGA GTGTGACTGT

 +1  V  C  R  G  S  T  G  G  S  R  G  P  F  E  G  K  P  I  P  N
1021 GTGTGCAGAG GGAGCACAGG AGGATCTAGA GGGCCCTTCG AAGGTAAGCC TATCCCTAAC

 +1  P  L  L  G  L  D  S  T  R  T  G  H  H  H  H  H  H
1081 CCTCTCCTCG GTCTCGATTC TACGCGTACC GGTCATCATC ACCATCACCA TTGA
```

FIG. 20. DNA and polypeptide sequence used for baculovirus/insect cell expression

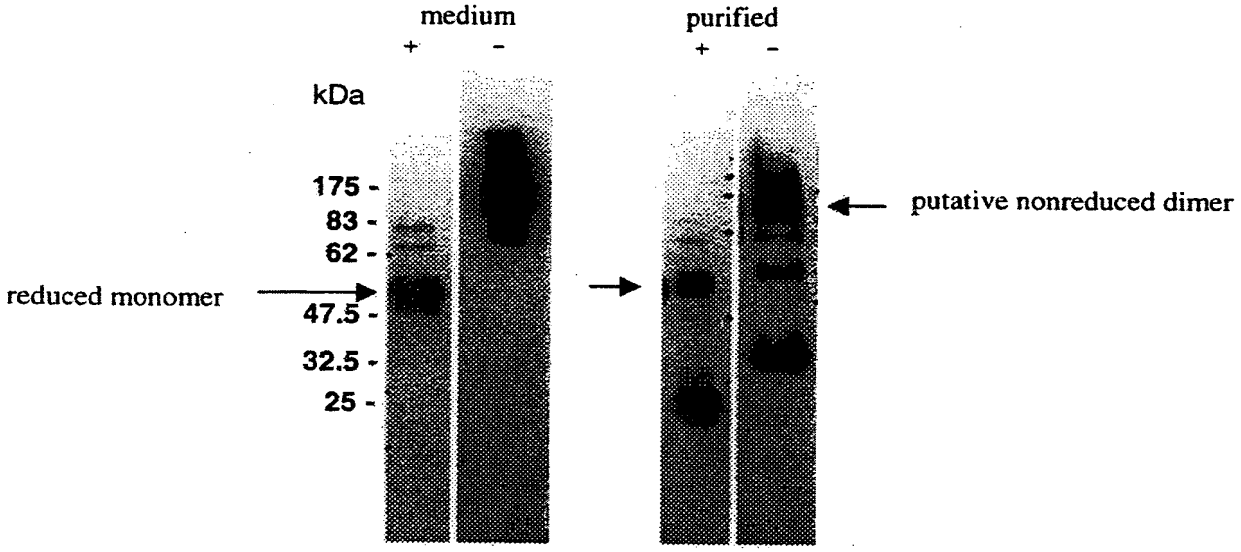
```
   1  GAATTCAAAG GCCTGTATTT TACTGTTTTC GTAACAGTTT TGTAATAAAA AAACCTATAA

  +3     m  k  f  l  v  n  v  a  l  v  f  m  v  v  y  i  s  y  i
  61  ATATGAAATT CTTAGTCAAC GTTGCCCTTG TTTTTATGGT CGTATACATT TCTTACATCT

  +3  y  a  D  P  E  S  H  H  H  H  H  H  E  S  N  L  S  S  K  F
 121  ATGCGGATCC GGAGTCTCAC CATCACCACC ATCATGAATC CAACCTGAGT AGTAAATTCC

  +3  Q  F  S  S  N  K  E  Q  N  G  V  Q  D  P  Q  H  E  R  I  I
 181  AGTTTTCCAG CAACAAGGAA CAGAACGGAG TACAAGATCC TCAGCATGAG AGAATTATTA

  +3  T  V  S  T  N  G  S  I  H  S  P  R  F  P  H  T  Y  P  R  N
 241  CTGTGTCTAC TAATGGAAGT ATTCACAGCC CAAGGTTTCC TCATACTTAT CCAAGAAATA

  +3  T  V  L  V  W  R  L  V  A  V  E  E  N  V  W  I  Q  L  T  F
 301  CGGTCTTGGT ATGGAGATTA GTAGCAGTAG AGGAAAATGT ATGGATACAA CTTACGTTTG

  +3  D  E  R  F  G  L  E  D  P  E  D  D  I  C  K  Y  D  F  V  E
 361  ATGAAAGATT TGGGCTTGAA GACCCAGAAG ATGACATATG CAAGTATGAT TTTGTAGAAG

  +3  V  E  E  P  S  D  G  T  I  L  G  R  W  C  G  S  G  T  V  P
 421  TTGAGGAACC CAGTGATGGA ACTATATTAG GGCGCTGGTG TGGTTCTGGT ACTGTACCAG

  +3  G  K  Q  I  S  K  G  N  Q  I  R  I  R  F  V  S  D  E  Y  F
 481  GAAAACAGAT TTCTAAAGGA AATCAAATTA GGATAAGATT TGTATCTGAT GAATATTTTC

  +3  P  S  E  P  G  F  C  I  H  Y  N  I  V  M  P  Q  F  T  E  A
 541  CTTCTGAACC AGGGTTCTGC ATCCACTACA ACATTGTCAT GCCACAATTC ACAGAAGCTG

  +3  V  S  P  S  V  L  P  P  S  A  L  P  L  D  L  L  N  N  A  I
 601  TGAGTCCTTC AGTGCTACCC CCTTCAGCTT TGCCACTGGA CCTGCTTAAT AATGCTATAA

  +3  T  A  F  S  T  L  E  D  L  I  R  Y  L  E  P  E  R  W  Q  L
 661  CTGCCTTTAG TACCTTGGAA GACCTTATTC GATATCTTGA ACCAGAGAGA TGGCAGTTGG

  +3  D  L  E  D  L  Y  R  P  T  W  Q  L  L  G  K  A  F  V  F  G
 721  ACTTAGAAGA TCTATATAGG CCAACTTGGC AACTTCTTGG CAAGGCTTTT GTTTTTGGAA

  +3  R  K  S  R  V  V  D  L  N  L  L  T  E  E  V  R  L  Y  S  C
 781  GAAAATCCAG AGTGGTGGAT CTGAACCTTC TAACAGAGGA GGTAAGATTA TACAGCTGCA

  +3  T  P  R  N  F  S  V  S  I  R  E  E  L  K  R  T  D  T  I  F
 841  CACCTCGTAA CTTCTCAGTG TCCATAAGGG AAGAACTAAA GAGAACCGAT ACCATTTTCT

  +3  W  P  G  C  L  L  V  K  R  C  G  G  N  C  A  C  C  L  H  N
 901  GGCCAGGTTG TCTCCTGGTT AAACGCTGTG GTGGGAACTG TGCCTGTTGT CTCCACAATT

  +3  C  N  E  C  Q  C  V  P  S  K  V  T  K  K  Y  H  E  V  L  Q
 961  GCAATGAATG TCAATGTGTC CCAAGCAAAG TTACTAAAAA ATACCACGAG GTCCTTCAGT

  +3  L  R  P  K  T  G  V  R  G  L  H  K  S  L  T  D  V  A  L  E
1021  TGAGACCAAA GACCGGTGTC AGGGGATTGC ACAAATCACT CACCGACGTG GCCCTGGAGC

  +3  H  H  E  E  S  D  C  V  C  R  G  S  T  G  G
1081  ACCATGAGGA GTGTGACTGT GTGTGCAGAG GGAGCACAGG AGGATAGCTC TAGA
```

FIG. 21. DNA and polypeptide sequence used for *E.coli* expression

```
  +3    Q  T  N  S  S  S  N  N  N  N  N  N  N  N  N  L  G  I
  1  CGCAGACTAA TTCGAGCTCG AACAACAACA ACAATAACAA TAACAACAAC CTCGGGATCG

  +3 E  G  R  I  S  E  F  E  S  N  L  S  S  K  F  Q  F  S  S  N
 61  AGGGAAGGAT TTCAGAATTC GAATCCAACC TGAGTAGTAA ATTCCAGTTT TCCAGCAACA

  +3 K  E  Q  N  G  V  Q  D  P  Q  H  E  R  I  I  T  V  S  T  N
121  AGGAACAGAA CGGAGTACAA GATCCTCAGC ATGAGAGAAT TATTACTGTG TCTACTAATG

  +3 G  S  I  H  S  P  R  F  P  H  T  Y  P  R  N  T  V  L  V  W
181  GAAGTATTCA CAGCCCAAGG TTTCCTCATA CTTATCCAAG AAATACGGTC TTGGTATGGA

  +3 R  L  V  A  V  E  E  N  V  W  I  Q  L  T  F  D  E  R  F  G
241  GATTAGTAGC AGTAGAGGAA AATGTATGGA TACAACTTAC GTTTGATGAA AGATTTGGGC

  +3 L  E  D  P  E  D  D  I  C  K  Y  D  F  V  E  V  E  E  P  S
301  TTGAAGACCC AGAAGATGAC ATATGCAAGT ATGATTTTGT AGAAGTTGAG GAACCCAGTG

  +3 D  G  T  I  L  G  R  W  C  G  S  G  T  V  P  G  K  Q  I  S
361  ATGGAACTAT ATTAGGGCGC TGGTGTGGTT CTGGTACTGT ACCAGGAAAA CAGATTTCTA

  +3 K  G  N  Q  I  R  I  R  F  V  S  D  E  Y  F  P  S  E  P  G
421  AAGGAAATCA AATTAGGATA AGATTTGTAT CTGATGAATA TTTTCCTTCT GAACCAGGGT

  +3 F  C  I  H  Y  N  I  V  M  P  Q  F  T  E  A  V  S  P  S  V
481  TCTGCATCCA CTACAACATT GTCATGCCAC AATTCACAGA AGCTGTGAGT CCTTCAGTGC

  +3 L  P  P  S  A  L  P  L  D  L  L  N  N  A  I  T  A  F  S  T
541  TACCCCCTTC AGCTTTGCCA CTGGACCTGC TTAATAATGC TATAACTGCC TTTAGTACCT

  +3 L  E  D  L  I  R  Y  L  E  P  E  R  W  Q  L  D  L  E  D  L
601  TGGAAGACCT TATTCGATAT CTTGAACCAG AGAGATGGCA GTTGGACTTA GAAGATCTAT

  +3 Y  R  P  T  W  Q  L  L  G  K  A  F  V  F  G  R  K  S  R  V
661  ATAGGCCAAC TTGGCAACTT CTTGGCAAGG CTTTTGTTTT TGGAAGAAAA TCCAGAGTGG

  +3 V  D  L  N  L  L  T  E  E  V  R  L  Y  S  C  T  P  R  N  F
721  TGGATCTGAA CCTTCTAACA GAGGAGGTAA GATTATACAG CTGCACACCT CGTAACTTCT

  +3 S  V  S  I  R  E  E  L  K  R  T  D  T  I  F  W  P  G  C  L
781  CAGTGTCCAT AAGGGAAGAA CTAAAGAGAA CCGATACCAT TTTCTGGCCA GGTTGTCTCC

  +3 L  V  K  R  C  G  G  N  C  A  C  C  L  H  N  C  N  E  C  Q
841  TGGTTAAACG CTGTGGTGGG AACTGTGCCT GTTGTCTCCA CAATTGCAAT GAATGTCAAT

  +3 C  V  P  S  K  V  T  K  K  Y  H  E  V  L  Q  L  R  P  K  T
901  GTGTCCCAAG CAAAGTTACT AAAAAATACC ACGAGGTCCT TCAGTTGAGA CCAAAGACCG

  +3 G  V  R  G  L  H  K  S  L  T  D  V  A  L  E  H  H  E  E  C
961  GTGTCAGGGG ATTGCACAAA TCACTCACCG ACGTGGCCCT GGAGCACCAT GAGGAGTGTG

  +3 D  C  V  C  R  G  S  T  G  G  H  H  H  H  H  H  *
1021 ACTGTGTGTG CAGAGGGAGC ACAGGAGGAC ATCATCACCA TCACCATTGA TCTAGAGTCG

1081 ACCTGCAGGC AAGCTT
```

FIG. 22. **Disulphide-linked dimerisation of VEGF-X**

(A) Mammalian cell expression



(B) *E.coli* expression

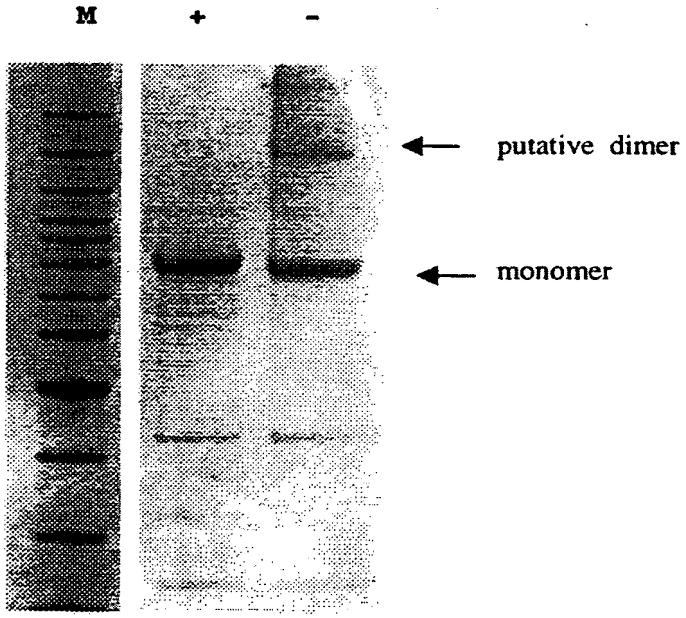


FIG. 23. **Glycosylation of VEGF-X**

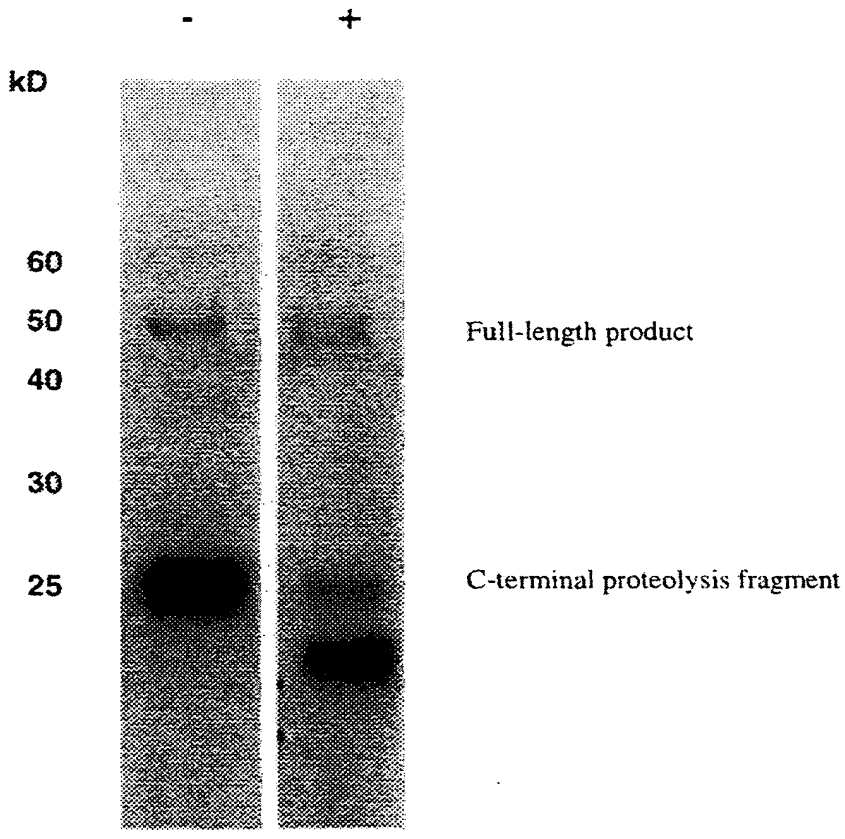


FIG. 24.

DNA and polypeptide sequence used for *E.coli* expression of the PDGF-like domain

```
  +3                  M  R  G  S  H  H  H  H  H  H  G  M  A  S  M
   1  AAGGAGATAT ACATATGCGG GGTTCTCATC ATCATCATCA TCATGGTATG GCTAGCATGA

  +3  T  G  G  Q  Q  M  G  R  D  L  Y  D  D  D  D  K  D  P  G  R
  61  CTGGTGGACA GCAAATGGGT CGGGATCTGT ACGACGATGA CGATAAGGAT CCGGGAAGAA

  +3  K  S  R  V  V  D  L  N  L  L  T  E  E  V  R  L  Y  S  C  T
 121  AATCCAGAGT GGTGGATCTG AACCTTCTAA CAGAGGAGGT AAGATTATAC AGCTGCACAC

  +3  P  R  N  F  S  V  S  I  R  E  E  L  K  R  T  D  T  I  F  W
 181  CTCGTAACTT CTCAGTGTCC ATAAGGGAAG AACTAAAGAG AACCGATACC ATTTTCTGGC

  +3  P  G  C  L  L  V  K  R  C  G  G  N  C  A  C  C  L  H  N  C
 241  CAGGTTGTCT CCTGGTTAAA CGCTGTGGTG GGAACTGTGC CTGTTGTCTC CACAATTGCA

  +3  N  E  C  Q  C  V  P  S  K  V  T  K  K  Y  H  E  V  L  Q  L
 301  ATGAATGTCA ATGTGTCCCA AGCAAAGTTA CTAAAAAATA CCACGAGGTC CTTCAGTTGA

  +3  R  P  K  T  G  V  R  G  L  H  K  S  L  T  D  V  A  L  E  H
 361  GACCAAAGAC CGGTGTCAGG GGATTGCACA AATCACTCAC CGACGTGGCC CTGGAGCACC

  +3  H  E  E  C  D  C  V  C  R  G  S  T  G  G
 421  ATGAGGAGTG TGACTGTGTG TGCAGAGGGA GCACAGGAGG ATAATGAATT CGAAGCTTGA

 481  TCCGGCTGCT AACAAAGCCC
```

FIG. 25. Expression of PDGF domain in *E.coli*

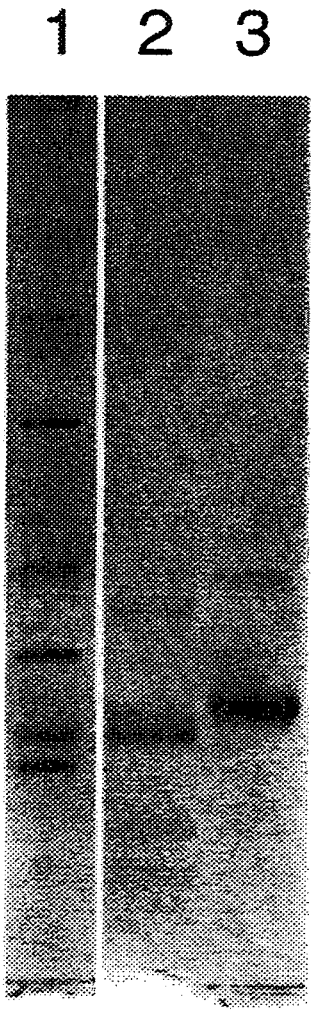


FIG. 26.

DNA and polypeptide sequence used for *E.coli* expression of the CUB-like domain

```
+2      M  A  M  D  I  G  I  N  S  D  P  E  S  H  H  H  H  H  H
1   GGCGATGGCC ATGGATATCG GAATTAATTC GGATCCGGAG TCTCACCATC ACCACCATCA

+2    E  S  N  L  S  S  K  F  Q  F  S  S  N  K  E  Q  N  G  V  Q
61  TGAATCCAAC CTGAGTAGTA AATTCCAGTT TTCCAGCAAC AAGGAACAGA ACGGAGTACA

+2    D  P  Q  H  E  R  I  I  T  V  S  T  N  G  S  I  H  S  P  R
121 AGATCCTCAG CATGAGAGAA TTATTACTGT GTCTACTAAT GGAAGTATTC ACAGCCCAAG

+2    F  P  H  T  Y  P  R  N  T  V  L  V  W  R  L  V  A  V  E  E
181 GTTTCCTCAT ACTTATCCAA GAAATACGGT CTTGGTATGG AGATTAGTAG CAGTAGAGGA

+2    N  V  W  I  Q  L  T  F  D  E  R  F  G  L  E  D  P  E  D  D
241 AAATGTATGG ATACAACTTA CGTTTGATGA AAGATTTGGG CTTGAAGACC CAGAAGATGA

+2    I  C  K  Y  D  F  V  E  V  E  E  P  S  D  G  T  I  L  G  R
301 CATATGCAAG TATGATTTTG TAGAAGTTGA GGAACCCAGT GATGGAACTA TATTAGGGCG

+2    W  C  G  S  G  T  V  P  G  K  Q  I  S  K  G  N  Q  I  R  I
361 CTGGTGTGGT TCTGGTACTG TACCAGGAAA ACAGATTTCT AAAGGAAATC AAATTAGGAT

+2    R  F  V  S  D  E  Y  F  P  S  E  P  G  F  C  I  H  Y  N  I
421 AAGATTTGTA TCTGATGAAT ATTTTCCTTC TGAACCAGGG TTCTGCATCC ACTACAACAT

+2    V  M  P  Q  F  T  E  A  V
481 TGTCATGCCA CAATTCACAG AAGCTGTGTA GTCGAGCTCC GTCGACAAGC TTGCGGCCGC

541 ACTCGAGCAC
```

FIG. 27. Expression of the CUB domain in *E.coli*

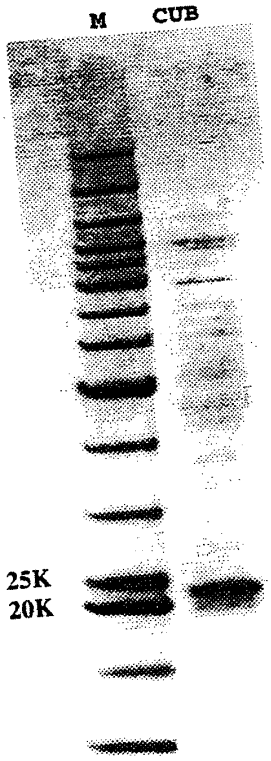


FIG. 28. The Effect of Truncated VEGF-X (CUB domain) on HUVEC Proliferation.

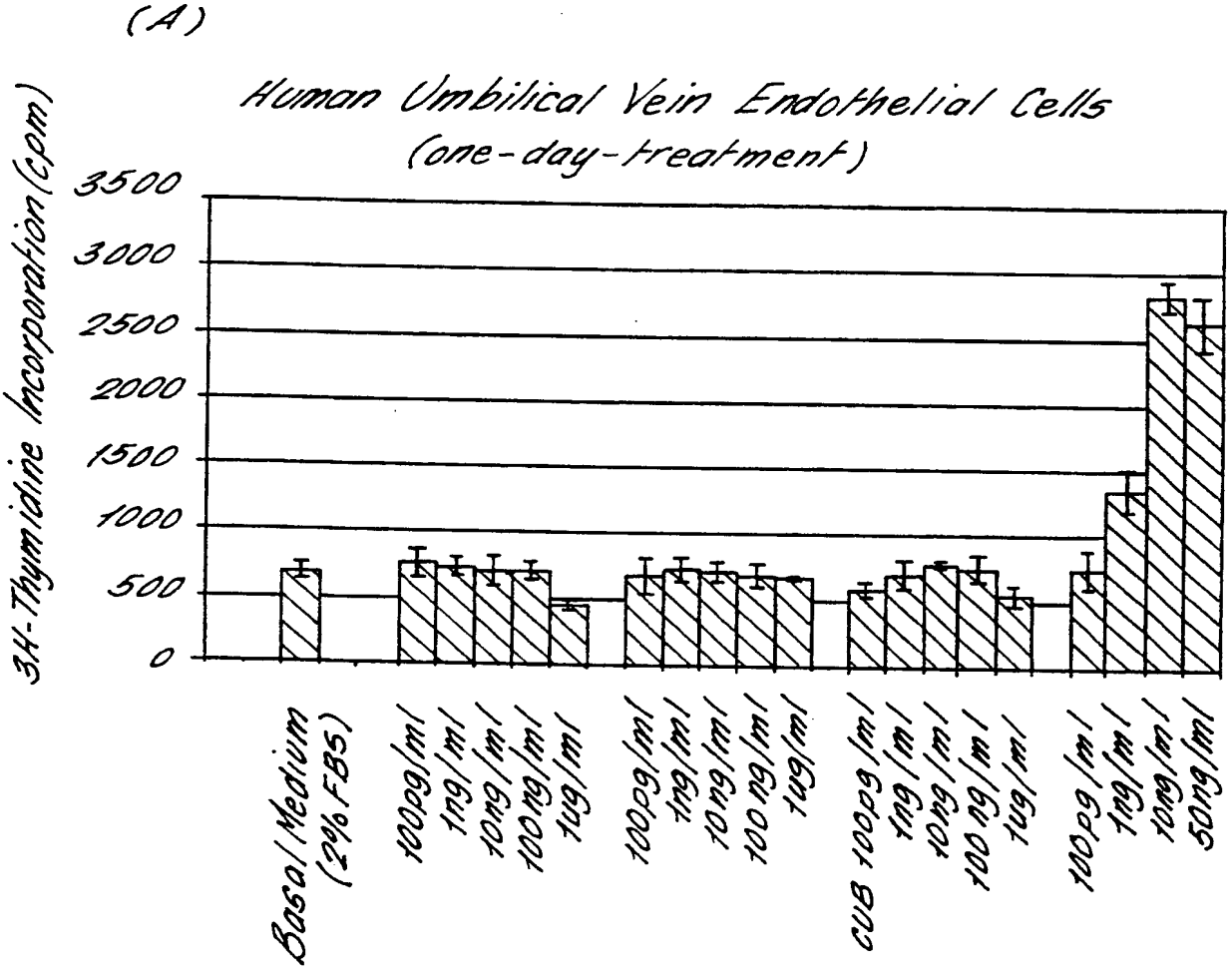


FIG. 28 (CONTINUED 1).

(B)

Human Umbilical Vein Endothelial Cells (24-hour-starving Followed by one-day-treatment)

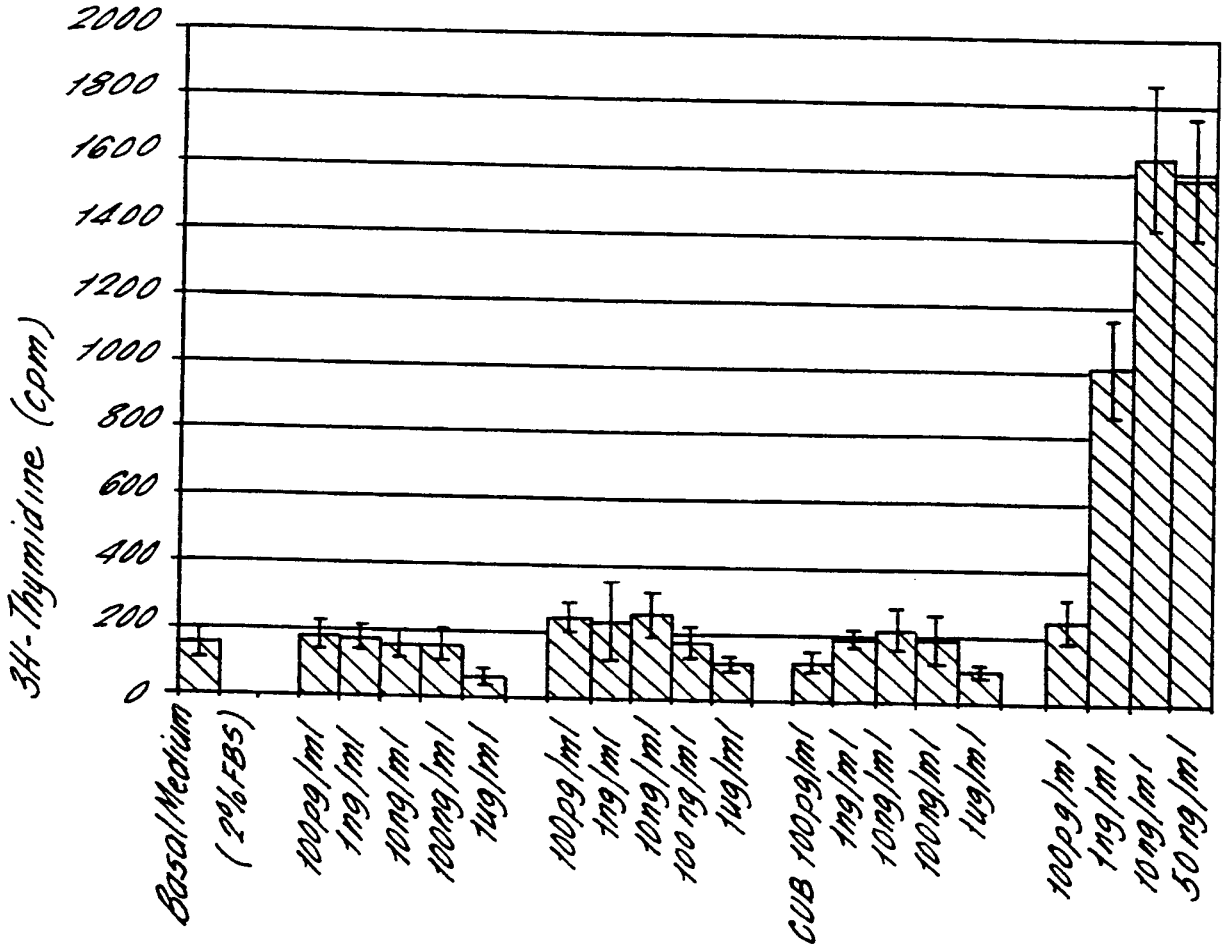


FIG. 28 (CONTINUED 2).

(C)

The effect of VEGF-A_{165} and VEGF-X CUB domain on the proliferation of HUVEC (two-day-treatment).

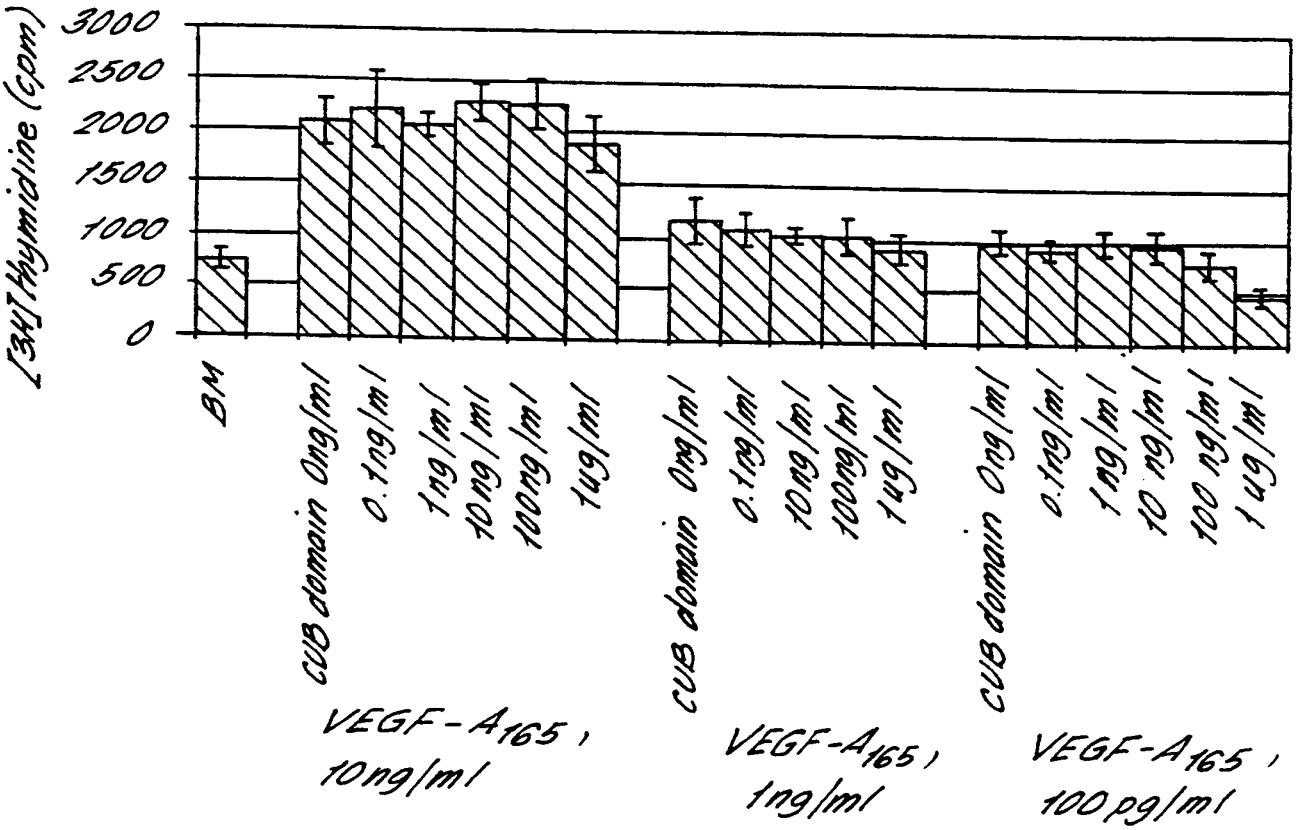
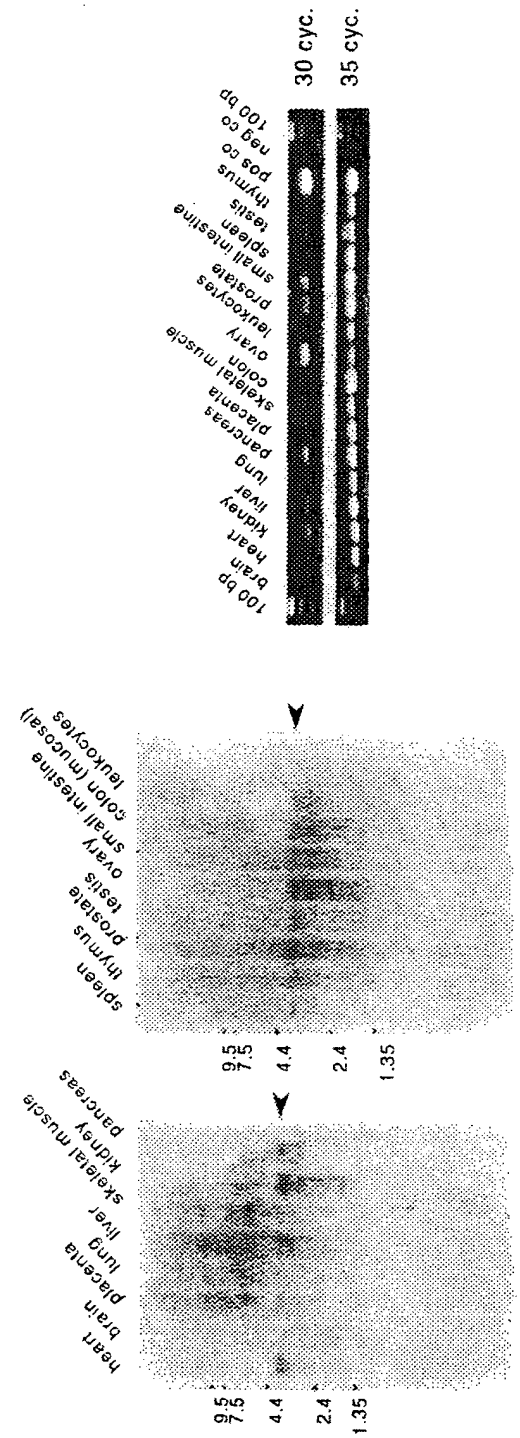


FIG. 29.

Tissue distribution of mRNA

(A) - Normal tissues



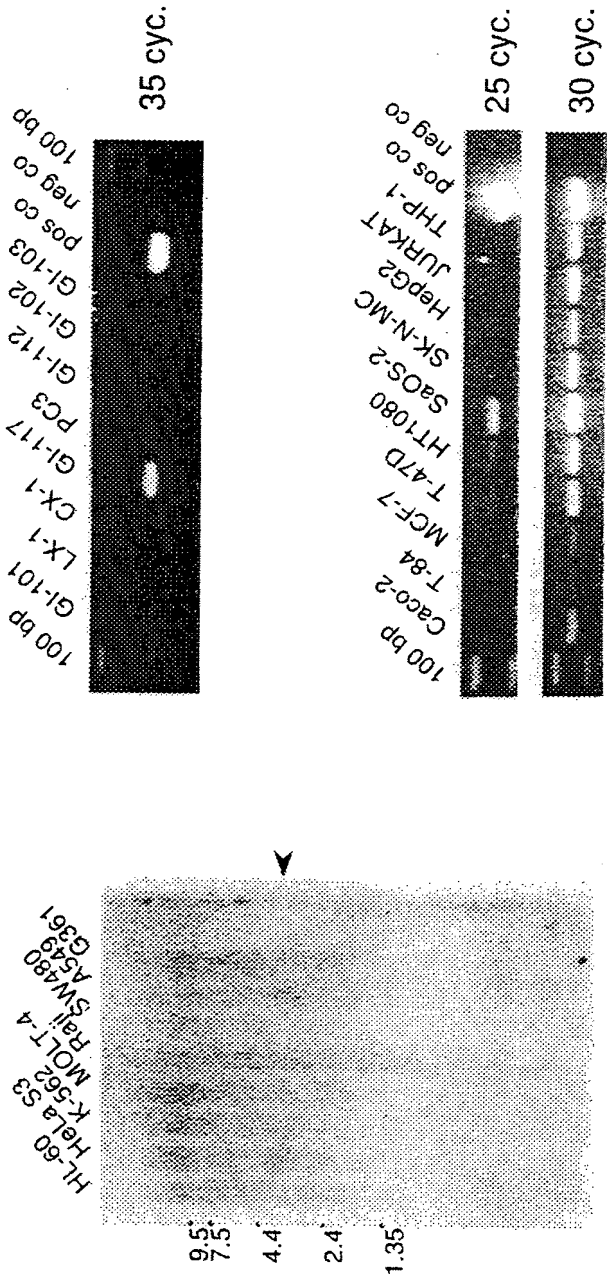


FIG 29 (CONTINUED). (B)- Tumour tissue and cell lines

FIG. 30.

Partial intron/exon structure of the VEGF-X gene

(A) - Genomic DNA sequences of 2 exons determined by sequencing

tttcttttataccatatagtggtggatctgaaccagGGTTCTGCATCCACTACAACATTGTCATGCCACAATTCACAGAAGCTGTG
AGTCCTTCAGTGCTACCCCCTTCAGCTTTGCCACTGGACCTGCTTAATAATGCTATAACTGCCTTTAGTACCTTGGAAGACCTTAT
TCGATATCTTGAACCAGAGAGATGGCAGTTGGACTTAGAAGATCTATATAGGCCAACTTGGCAACTTCTTGGCAAGGCTTTTGTTT
TTGGAAGAAAATCCAGAGTGGTGGATCTGAACCTTCTAACAGAGGAGGTAAGATTATACAGCTGCACACCTCGTAACTTCTCAGTG
TCCATAAGGGAAGAACTAAAGAGAACCGATACCATTTTCTGGCCAGGTTGTCTCCTGGTTAAACGCTGTGGTGGGAACTGTGCCTG
TTGTCTCCACAATTGCAATGAATGTCAATGTGTCCCAAGCAAAGTTACTAAAAAATACCACGAGgtaggtatacaatttttttttt
ggtttccttcgggtattttatgtctt

aaagccagtcatagacattcgttgatttttaaaagtggcttactcttattccctttcagGTCCTTCAGTTGAGACCAAAGACCGGT
GTCAGGGGATTGCACAAATCACTCACCGACGTGGCCCTGGAGCACCATGAGGAGTGTGACTGTGTGTGCAGAGGGAGCACAGGAGG
ATAGCCGCATCACCACCAGCAGCTCTTGCCCAGAGCTGTGCAGTGCAGTGGCTGATTCTATTAGAGAACGTATGCGTTATCTCCAT
CCTTAATCTCAGTTGTTTGCTTCAAGGACCTTTCATCTTCAGGATTTACAGTGCATTCTGAAAGAGGAGACATCAAACAGAATTAG
GAGTTGTGCAACAGCTCTTTTGAGAGGAGGCCTAAAGGACAGGAGAAAAGGTCTTCAATCGTGGAAAGAAAATTAAATGTTGTATT
AAATAGATCACCAGCTAGTTTCAGAGTTACCATGTACGTATTCCACTAGCTGGGTTCTGTATTTCAGTTCTTTCGATACGGCTTAG
GGTAATGTCAGTACAGGAAAAAAACTGTGCAAGTGAGCACCTGATTCCGTTGCCTTGGCTTAACTCTAAAGCTCCATGTCCTGGGC
CTAAAATCGTATAAAATCTGGATTTTTTTTTTTTTTTTGCGCATATTCACATATGTAAACCAGAACATTCTATGTACTACAAACC
TGGTTTTTAAAAAGGAACTATGTTGCTATGAATTAAACTTGTGTCATGCTGATAGGACAGACTGGATTTTTCATATTTCTTATTAA
AATTTCTGCCATTTAGAAGAAGAGAACTACATTCATGGTTTGGAAGAGATAAACCTGAAAAGAAGAGTGGCCTTATCTTCACTTTA
TCGATAAGTCAGTTTATTTGTTTCATTGTGTACATTTTTATATTCTCCTTTTGACATTATAACTGTTGGCTTTTCTAATCTTGTTA
AATATATCTATTTTTTACCAAAGGTATTTAATATTCTTTTTTATGACAACTTAGATCAACTATTTTTAGCTTGGTAAATTTTTCTAA
ACACAATTGTTATAGCCAGAGGAACAAAGATGATATAAAATATTGTTGCTCTGACAAAAATACATGTATTTCATTCTCGTATGGTG
CTAGAGTTAGATTAATCTGCATTTTAAAAAACTGAATTGGAATAGAATTGGTAAGTTGCAAAGACTTTTTGAAAATAATTAAATTA
TCATATCTTCCATTCCTGTTATTGGAGATGAAAATAAAAAGCAACTTATGAAAGTAGACATTCAGATCCAGCCATTACTAACCTAT
TCCTTTTTTGGGGAAATCTGAGCCTAGCTCAGAAAAACATAAAGCACCTTGAAAAAGACTTGGCAGCTTCCTGATAAAGCGTGCTG
TGCTGTGCAGTAGGAACACATCCTATTTATTGTGATGTTGTGGTTTTATTATCTTAAACTCTGTTCCATACACTTGTATAAATACA
TGGATATTTTTATGTACAGAAGTATGTCTCTTAACCAGTTCACTTATTGTACTCTGGCAATTTAAAAGAAAATCAGTAAAATATTT
TGCTTGTAAAATGCTTAATATCGTGCCTAGGTTATGTGGTGACTATTTGAATCAAAAATGTATTGAATCATCAAATAAAAGAATGT
GGCTATTTTGGGGAGAAAATTatgtgtgtgtgtgctcaagatttatttcttggactctgagaaaatgaaagataaa

FIG. 30 (CONTINUED 1).

(B) - Location of splice sites within the cDNA sequence

```
   1  GAATTCGCCC TTTTGTTTAA ACCTTGGGAA CTGGTTCAGG TCCAGGTTTT GCTTTGATCC

  61  TTTTCAAAAA CTGGAGACAC AGAAGAGGGC TCTAGGAAAA AGTTTTGGAT GGGATTATGT

 121  GGAAACTACC CTGCGATTCT CTGCTGCCAG AGCAGGCTCG GCGCTTCCAC CCCAGTGCAG

 181  CCTTCCCCTG GCGGTGGTGA AAGAGACTCG GGAGTCGCTG CTTCCAAAGT GCCCGCCGTG

  +3                                M  S  L  F  G  L  L  L  L  T  S
 241  AGTGAGCTCT CACCCCAGTC AGCCAAATGA GCCTCTTCGG GCTTCTCCTG CTGACATCTG

  +3 A  L  A  G  Q  R  Q  G  T  Q  A  E  S  N  L  S  S  K  F  Q
 301  CCCTGGCCGG CCAGAGACAG GGGACTCAGG CGGAATCCAA CCTGAGTAGT AAATTCCAGT

  +3 F  S  S  N  K  E  Q  N  G  V  Q  D  P  Q  H  E  R  I  I  T
 361  TTTCCAGCAA CAAGGAACAG AACGGAGTAC AAGATCCTCA GCATGAGAGA ATTATTACTG

  +3 V  S  T  N  G  S  I  H  S  P  R  F  P  H  T  Y  P  R  N  T
 421  TGTCTACTAA TGGAAGTATT CACAGCCCAA GGTTTCCTCA TACTTATCCA AGAAATACGG

  +3 V  L  V  W  R  L  V  A  V  E  E  N  V  W  I  Q  L  T  F  D
 481  TCTTGGTATG GAGATTAGTA GCAGTAGAGG AAAATGTATG GATACAACTT ACGTTTGATG

  +3 E  R  F  G  L  E  D  P  E  D  D  I  C  K  Y  D  F  V  E  V
 541  AAAGATTTGG GCTTGAAGAC CCAGAAGATG ACATATGCAA GTATGATTTT GTAGAAGTTG

  +3 E  E  P  S  D  G  T  I  L  G  R  W  C  G  S  G  T  V  P  G
 601  AGGAACCCAG TGATGGAACT ATATTAGGGC GCTGGTGTGG TTCTGGTACT GTACCAGGAA

  +3 K  Q  I  S  K  G  N  Q  I  R  I  R  F  V  S  D  E  Y  F  P
 661  AACAGATTTC TAAAGGAAAT CAAATTAGGA TAAGATTTGT ATCTGATGAA TATTTTCCTT

  +3 S  E  P |G  F  C  I  H  Y  N  I  V  M  P  Q  F  T  E  A  V
 721  CTGAACCAG|G GTTCTGCATC CACTACAACA TTGTCATGCC ACAATTCACA GAAGCTGTGA

  +3 S  P  S  V  L  P  P  S  A  L  P  L  D  L  L  N  N  A  I  T
 781  GTCCTTCAGT GCTACCCCCT TCAGCTTTGC CACTGGACCT GCTTAATAAT GCTATAACTG

  +3 A  F  S  T  L  E  D  L  I  R  Y  L  E  P  E  R  W  Q  L  D
 841  CCTTTAGTAC CTTGGAAGAC CTTATTCGAT ATCTTGAACC AGAGAGATGG CAGTTGGACT

  +3 L  E  D  L  Y  R  P  T  W  Q  L  L  G  K  A  F  V  F  G  R
 901  TAGAAGATCT ATATAGGCCA ACTTGGCAAC TTCTTGGCAA GGCTTTTGTT TTTGGAAGAA

  +3 K  S  R  V  V  D  L  N  L  L  T  E  E  V  R  L  Y  S  C  T
 961  AATCCAGAGT GGTGGATCTG AACCTTCTAA CAGAGGAGGT AAGATTATAC AGCTGCACAC

  +3 P  R  N  F  S  V  S  I  R  E  E  L  K  R  T  D  T  I  F  W
1021  CTCGTAACTT CTCAGTGTCC ATAAGGGAAG AACTAAAGAG AACCGATACC ATTTTCTGGC

  +3 P  G  C  L  L  V  K  R  C  G  G  N  C  A  C  C  L  H  N  C
1081  CAGGTTGTCT CCTGGTTAAA CGCTGTGGTG GGAACTGTGC CTGTTGTCTC CACAATTGCA

  +3 N  E  C  Q  C  V  P  S  K  V  T  K  K  Y  H  E |V  L  Q  L
1141  ATGAATGTCA ATGTGTCCCA AGCAAAGTTA CTAAAAAATA CCACGAG|GTC CTTCAGTTGA
```

FIG. 30 (CONTINUED 2).

```
   +3 R  P  K  T  G  V  R  G  L  H  K  S  L  T  D  V  A  L  E  H
1201  GACCAAAGAC CGGTGTCAGG GGATTGCACA AATCACTCAC CGACGTGGCC CTGGAGCACC

   +3 H  E  E  C  D  C  V  C  R  G  S  T  G  G
1261  ATGAGGAGTG TGACTGTGTG TGCAGAGGGA GCACAGGAGG ATAGCCGCAT CACCACCAGC

1321  AGCTCTTGCC CAGAGCTGTG CAGTGCAGTG GCTGATTCTA TTAGAGAACG TATGCGTTAT

1381  CTCCATCCTT AATCTCAGTT GTTTGCTTCA AGGACCTTTC ATCTTCAGGA TTTACAGTGC

1441  ATTCTGAAAG AGGAGACATC AAACAGAATT AGGAGTTGTG CAACAGCTCT TTTGAGAGGA

1501  GGCCTAAAGG ACAGGAGAAA AGGTCTTCAA TCGTGGAAAG AAAATTAAAT GTTGTATTAA

1561  ATAGATCACC AGCTAGTTTC AGAGTTACCA TGTACGTATT CCACTAGCTG GGTTCTGTAT

1621  TTCAGTTCTT TCGATACGGC TTAGGGTAAT GTCAGTACAG GAAAAAAACT GTGCAAGTGA

1681  GCACCTGATT CCGTTGCCTT GCTTAACTCT AAAGCTCCAT GTCCTGGGCC TAAAATCGTA

1741  TAAAATCTGG ATTTTTTTTT TTTTTTTTG CTCATATTCA CATATGTAAA CCAGAACATT

1801  CTATGTACTA CAAACCTGGT TTTTAAAAAG GAACTATGTT GCTATGAATT AAACTTGTGT

1861  CATGCTGATA GGACAGACTG GATTTTTCAT ATTTCTTATT AAAATTTCTG CCATTTAGAA

1921  GAAGAGAACT ACATTCATGG TTTGGAAGAG ATAAACCTGA AAAGAAGAGT GGCCTTATCT

1981  TCACTTTATC GATAAGTCAG TTTATTTGTT TCATTGTGTA CATTTTTATA TTCTCCTTTT

2041  GACATTATAA CTGTTGGCTT TTCTAATCTT GTTAAATATA TCTATTTTTA CCAAAGGTAT

2101  TTAATATTCT TTTTTATGAC AACTTAGATC AACTATTTTT AGCTTGGTAA ATTTTTCTAA

2161  ACACAATTGT TATAGCCAGA GGAACAAAGA TGATATAAAA TATTGTTGCT CTGACAAAAA

2221  TACATGTATT TCATTCTCGT ATGGTGCTAG AGTTAGATTA ATCTGCATTT TAAAAAACTG

2281  AATTGGAATA GAATTGGTAA GTTGCAAAGA CTTTTTGAAA ATAATTAAAT TATCATATCT

2341  TCCATTCCTG TTATTGGAGA TGAAAATAAA AAGCAACTTA TGAAAGTAGA CATTCAGATC

2401  CAGCCATTAC TAACCTATTC CTTTTTTGGG GAAATCTGAG CCTAGCTCAG AAAAACATAA

2461  AGCACCTTGA AAAAGACTTG GCAGCTTCCT GATAAAGCGT GCTGTGCTGT GCAGTAGGAA

2521  CACATCCTAT TTATTGTGAT GTTGTGGTTT TATTATCTTA AACTCTGTTC CATACACTTG

2581  TATAAATACA TGGATATTTT TATGTACAGA AGTATGTCTC TTAACCAGTT CACTTATTGT

2641  ACCTGGAAGG GCGAATTCTG CAGATATC
```

FIG. 31.

The Effect of Fl-VEGF-X on HUVEC Proliferation:
(24-hour serum starvation Followed by
one day-treatment)

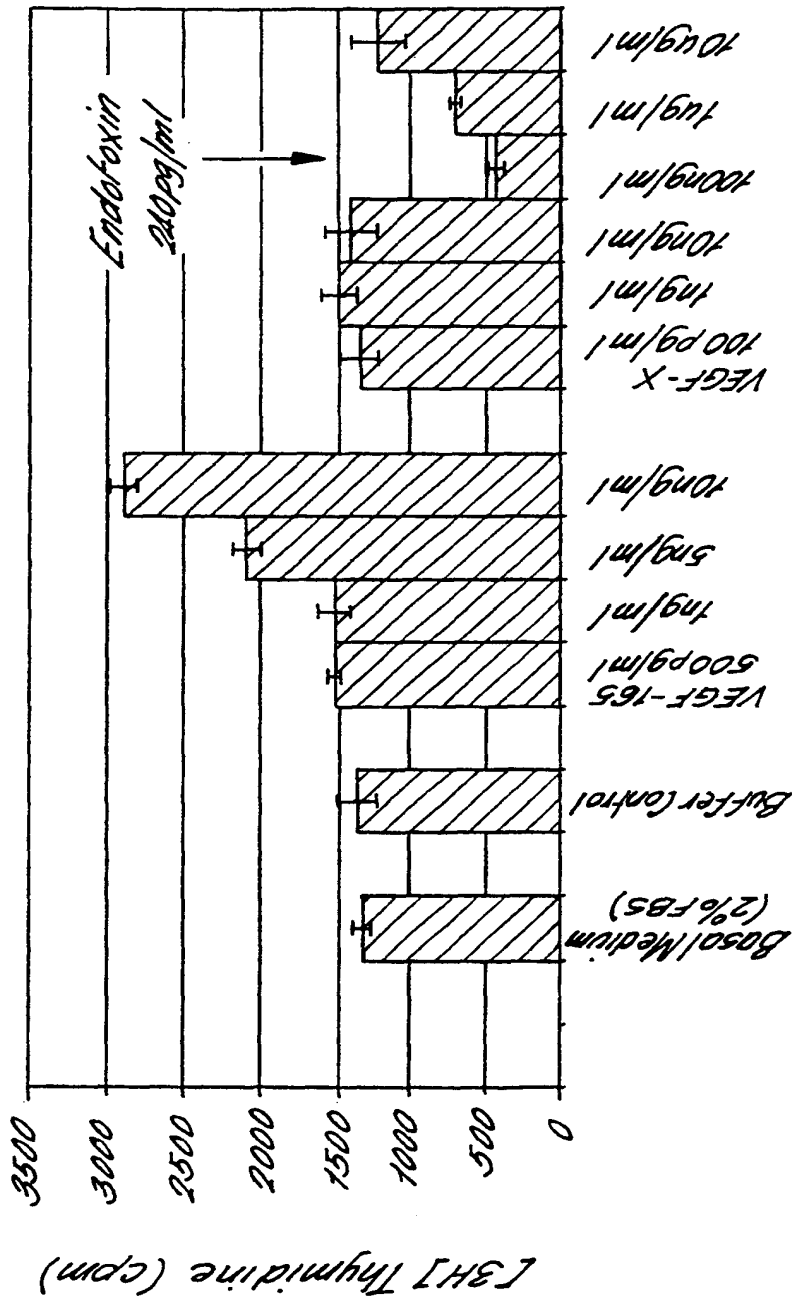


FIG. 32.

The Combined Effect of Truncated VEGF-X (CUB domain) and Human Recombinant $VEGF_{165}$ on HUVEC Proliferation: (24-hour serum starvation Followed by two-day-treatment)

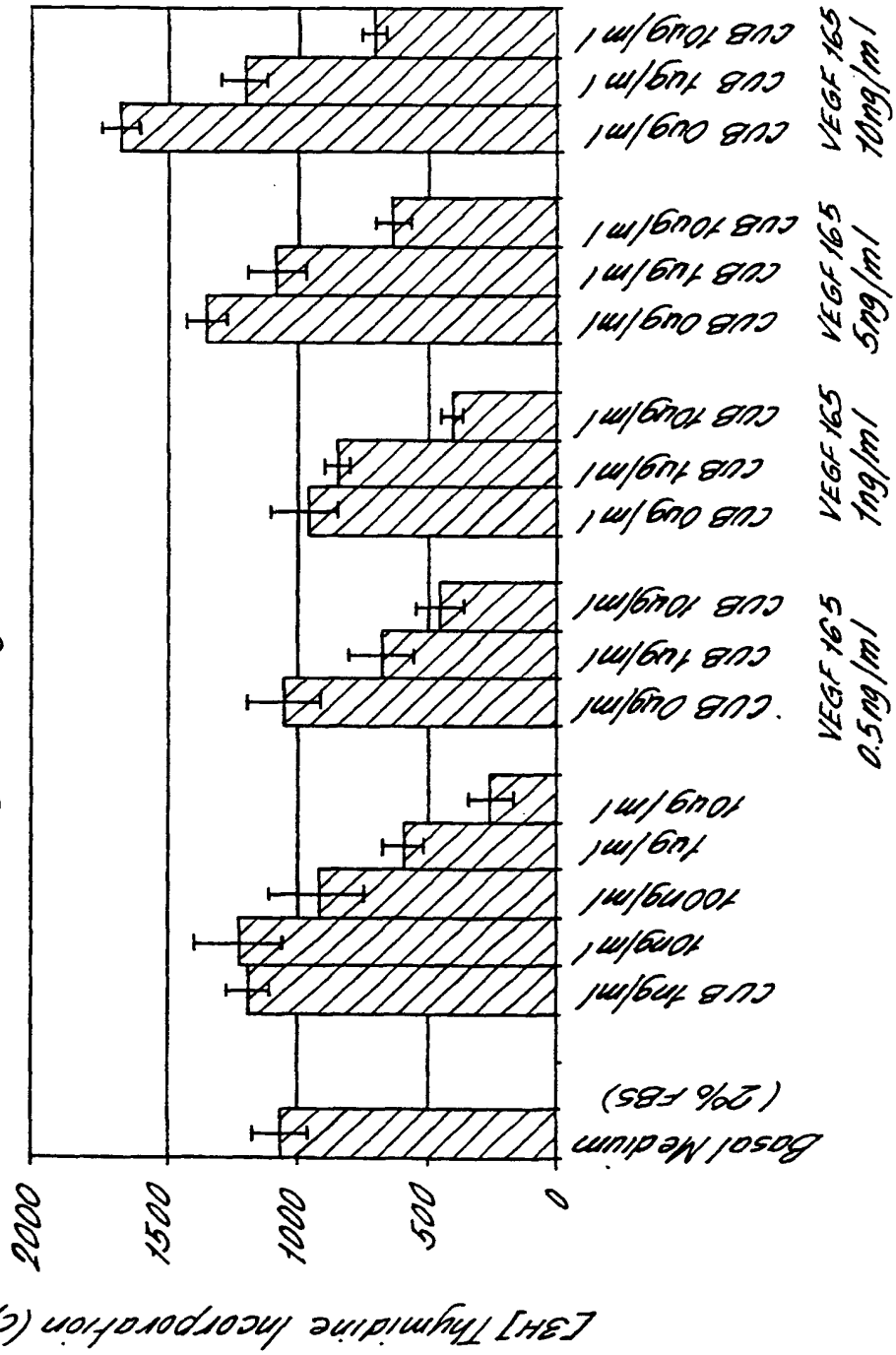
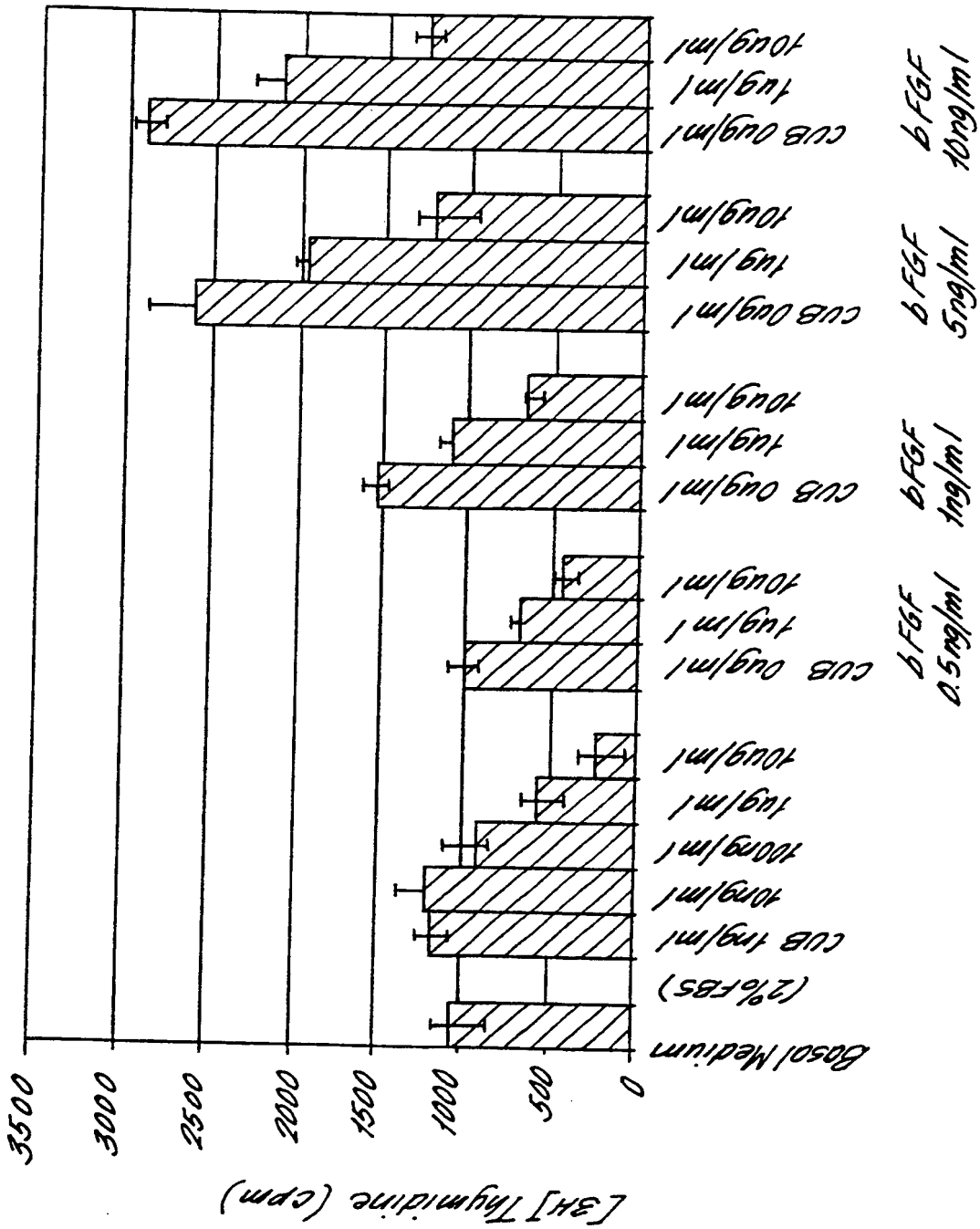


FIG. 33.
The Combined Effect of CUB Domain and Human Recombinant bFGF on HUVEC Proliferation: (24-hour serum starvation Followed by two-day-treatment).



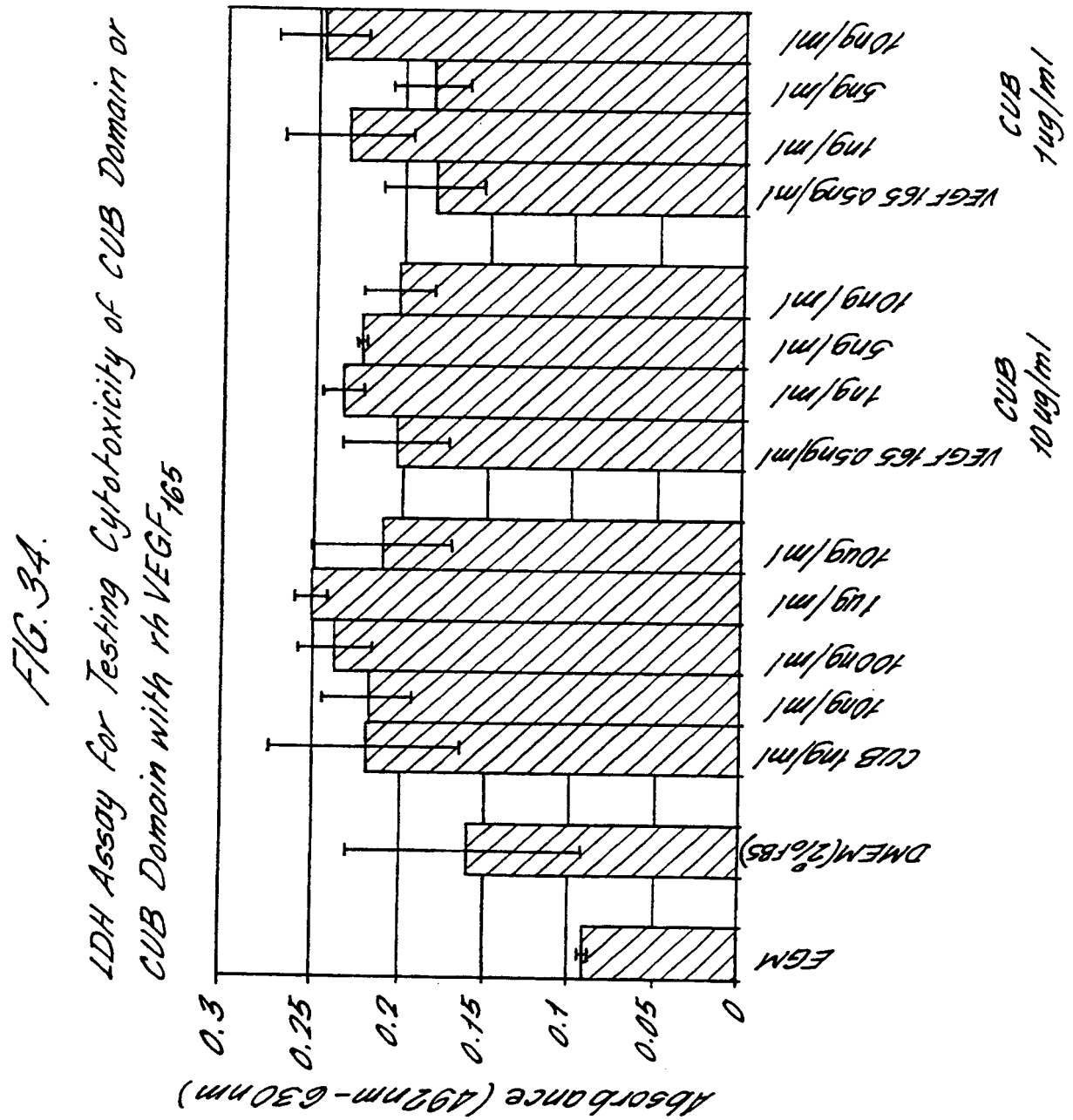
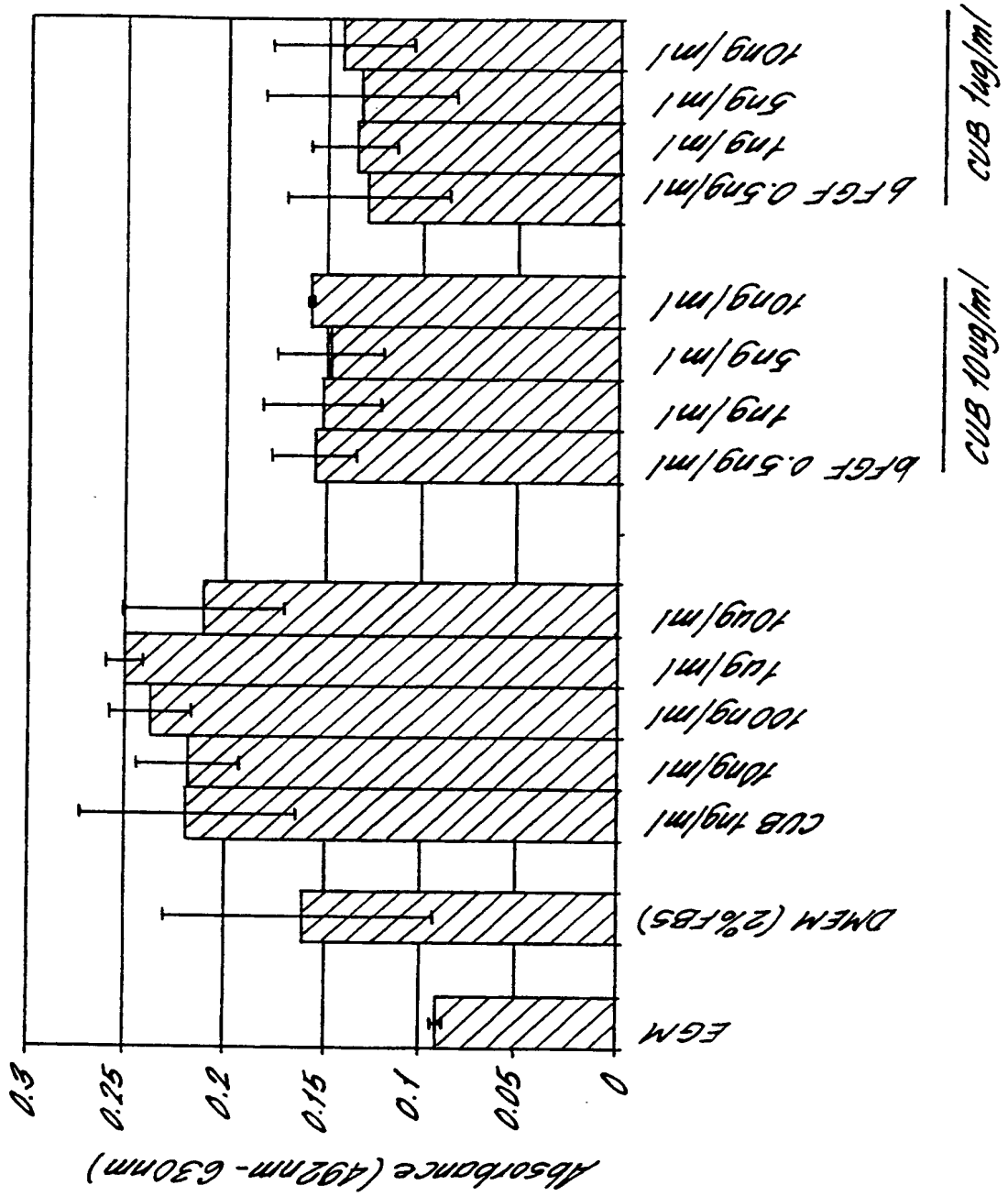


FIG. 34.

LDH Assay for Testing Cytotoxicity of CUB Domain or CUB Domain with rh $VEGF_{165}$

FIG. 35.

LDH Assay For Testing Cytotoxicity of CUB Domain or CUB Domain with rh-bFGF



Applicant's or agent's file reference	International application No.
B0192/7011WO	PCT/US99/30503

INDICATIONS RELATING TO DEPOSITED MICROORGANISM OR OTHER BIOLOGICAL MATERIAL

(PCT Rule 13*bis*)

A. The indications made below relate to the deposited microorganism or other biological material referred to in the description on page 21, line 15-16

B. IDENTIFICATION OF DEPOSIT Further deposits are identified on an additional sheet ☐

Name of depositary institution
BELGIAN COORDINATED COLLECTIONS OF MICROORGANISMS (BCCM)™
LABORATORIUM VOOR MOLECULAIRE BIOLOGIE - PLASMIDENCOLLECTIE (LMBP)

Address of depositary institution *(including postal code and country)*
Universiteit Gent
K.L. Ledeganckstraat 35
B-9000 Gent, Belgium

Date of deposit	Accession Number
20 December 1999 (20.12.99)	LMBP 3991

C. ADDITIONAL INDICATIONS *(leave blank if not applicable)* This information is continued on an additional sheet ☐

D. DESIGNATED STATES FOR WHICH INDICATIONS ARE MADE *(if the indications are not for all designated States)*

E. SEPARATE FURNISHING OF INDICATIONS *(leave blank if not applicable)*

The indications listed below will be submitted to the International Bureau later *(specify the general nature of the indications e.g., "Accession Number of Deposit")*

For receiving Office use only	For International Bureau use only
☐ This sheet was received with the international application	☒ This sheet was received by the International Bureau on: 19 APRIL 2000 (19.04.00)
Authorized officer	Authorized officer Ellen Moyse

BELGIAN COORDINATED COLLECTIONS OF MICROORGANISMS - BCCM™ LMBP-COLLECTION

Page 1 of Form BCCM™/LMBP/BP/4/99-23 Receipt in the case of an original deposit

---

Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure

Receipt in the case of an original deposit issued pursuant to Rule 7.1 by the International Depositary Authority BCCM™/LMBP identified at the bottom of next page

International **Form BCCM™/LMBP/BP/4/99-23**

To : **Name of the depositor :** Janssen Pharmaceutica N.V.

Address : Turnhoutseweg 30
B-2340 Beerse
Belgium

I. Identification of the microorganism:

I.1 Identification reference given by the **depositor:**

VEGF-X CUB PET22b

I.2 Accession number given by the **International Depositary Authority:**

LMBP 3991

BELGIAN COORDINATED COLLECTIONS OF MICROORGANISMS - BCCM™ LMBP-COLLECTION
Page 2 of Form BCCM™/LMBP/BP/4/99-23 Receipt in the case of an original deposit

---

II. Scientific description and/or proposed taxonomic designation

The microorganism identified under I above was accompanied by:

(mark with a cross the applicable box(es))

– a scientific description	yes ☒	no ☐
– a proposed taxonomic designation	yes ☐	no ☒

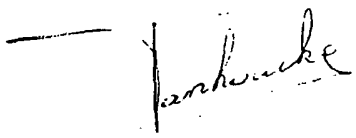
III. Receipt and acceptance

This International Depositary Authority accepts the microorganism identified under I above, which was received by it on (date of original deposit) : December 20, 1999

IV. International Depositary Authority

Belgian Coordinated Collections of Microorganisms (BCCM™)
Laboratorium voor Moleculaire Biologie - Plasmidencollectie (LMBP)
Universiteit Gent
K.L. Ledeganckstraat 35
B-9000 Gent, Belgium

Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorized official(s):



Date : January 12, 2000

Martine Vanhoucke
BCCM/LMBP curator

BELGIAN COORDINATED COLLECTIONS OF MICROORGANISMS - BCCM™
LMBP-COLLECTION
Page 1 of Form BCCM™/LMBP/BP/9/99-23 Viability statement

Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purposes of Patent Procedure

Viability statement issued pursuant to Rule 10.2 by the International Depositary Authority BCCM™/LMBP identified on the following page

International Form BCCM™/LMBP/BP/9/99-23

To : Party to whom the viability statement is issued:

Name : Dr Filip De Corte

Address : Janssen Pharmaceutica N.V.
Turnhoutseweg 30
B-2340 Beerse
Belgium

I. Depositor:

I.1 Name : Janssen Pharmaceutica N.V.

I.2 Address : Turnhoutseweg 30
B-2340 Beerse
Belgium

II. Identification of the microorganism:

II.1 Accession number given by the International Depositary Authority:

LMBP 3991

II.2 Date of the original deposit (or where a new deposit or a transfer has been made, the most recent relevant date) : December 20, 1999

III. Viability statement.

The viability of the microorganism identified under II above was tested on

: January 11, 2000

(Give date. In the cases referred to in Rule 10.2(a)(ii) and (iii), refer to the most recent viability test).

On that date, the said microorganism was: (mark the applicable box with a cross)

☒ viable

☐ no l nger viable

BELGIAN COORDINATED COLLECTIONS OF MICROORGANISMS

BELGIAN COORDINATED COLLECTIONS OF MICROORGANISMS - BCCM™
LMBP-COLLECTION
Page 2 of Form BCCM™/LMBP/BP/9/99-23 Viability statement

IV. **Conditions under which the viability test has been performed:**

(Fill in if the information has been requested and if the results of the test were negative).

V. **International Depositary Authority**

Belgian Coordinated Collections of Microorganisms (BCCM™)
Laboratorium voor Moleculaire Biologie - Plasmidencollectie (LMBP)
Universiteit Gent
K.L. Ledeganckstraat 35
B-9000 Gent, Belgium

Signature(s) of person(s) having the power to represent the International Depositary Authority or of authorized official(s):

Date : January 12, 2000

Martine Vanhoucke
BCCM/LMBP curator